


```
QY 61 GAGGCTCCAAAGCTTTGCTTCTCTGTGGCTGGTTCAGACGGCAGCGACACCTTCGTGG 120
DB |||||
QY 61 GAGGCTCCAAAGCTTTGCTTCTCTGTGGCTGGTTCAGACGGCAGCGACACCTTCGTGG 120
DB |||||
QY 121 CAGGAGTTGGGACGAATTCAGGACGCTCAAGAAGACCTCTCAAGAGACCTTCCCGGTGGA 180
DB |||||
QY 121 CAGGAGTTGGGACGAATTCAGGACGCTCAAGAAGACCTCTCAAGAGACCTTCCCGGTGGA 180
DB |||||
QY 181 GCGGGCTGTCTGCGGAGATCTGACCGGTTCTCCAAAGCTTCTCGATGCAACCACTGTT 240
DB |||||
QY 181 GCGGGCTGTCTGCGGAGATCTGACCGGTTCTCCAAAGCTTCTCGATGCAACCACTGTT 240
DB |||||
QY 241 GGGACGCTGGGGCCGACGAGCGCGGCTGCGCGCTGCGAGCTGTTGGAACTATTTC 300
DB |||||
QY 241 GGGACGCTGGGGCCGACGAGCGCGGCTGCGCGCTGCGAGCTGTTGGAACTATTTC 300
DB |||||
QY 301 TCGGAGGCTGTGCGGACTGACAGAGCGGTGGCACGGAGCCGACGATCACTGGGCTTCTT 360
DB |||||
QY 301 TCGGAGGCTGTGCGGACTGACAGAGCGGTGGCACGGAGCCGACGATCACTGGGCTTCTT 360
DB |||||
QY 361 CGCACCGCAACCCCTGGACCTGGAGCCCGCGCTGCCACCCGCGACCGGGTGATCCTGCC 420
DB |||||
QY 361 CGCACCGCAACCCCTGGACCTGGAGCCCGCGCTGCCACCCGCGACCGGGTGATCCTGCC 420
DB |||||
QY 421 CACCCAGAGGAGCAGCTTCTTCTGCGCTGCGGGCCGCTCTCCATCCACAGTCTGGA 480
DB |||||
QY 421 CACCCAGAGGAGCAGCTTCTTCTGCGCTGCGGGCCGCTCTCCATCCACAGTCTGGA 480
DB |||||
QY 481 GGCTCAGAGCTGCGCTGCTGACGCTTCTGTATCCAGGACACCGGGATAGGCCCTTT 540
DB |||||
QY 481 GGCTCAGAGCTGCGCTGCTGACGCTTCTGTATCCAGGACACCGGGATAGGCCCTTT 540
DB |||||
QY 541 TCAGGCGCAGGCCACAGGAGCTTGGACGTGCTGTGCGGCAACCCCTCAGGCTGTGGCT 600
DB |||||
QY 541 TCAGGCGCAGGCCACAGGAGCTTGGACGTGCTGTGCGGCAACCCCTCAGGCTGTGGCT 600
DB |||||
QY 601 GGTGAGAAAGAGACCGGACGACCGCTGTGTTTCCAGGCGCCCTACCTCGAGAGGCGGC 660
DB |||||
QY 601 GGTGAGAAAGAGACCGGACGACCGCTGTGTTTCCAGGCGCCCTACCTCGAGAGGCGGC 660
DB |||||
QY 661 CCCGGGCAAGGCGCGGAGGAGCGCTGCTTCCAGGAGACGCGCTGCCAGTCTGTGTC 720
DB |||||
QY 661 CCCGGGCAAGGCGCGGAGGAGCGCTGCTTCCAGGAGACGCGCTGCCAGTCTGTGTC 720
DB |||||
QY 721 TTCCGCGCCTACGAGAGACCGCGCAGATGAGCTGTCCGTCGCCGCGGGGCGCGGT 780
DB |||||
QY 721 TTCCGCGCCTACGAGAGACCGCGCAGATGAGCTGTCCGTCGCCGCGGGGCGCGGT 780
DB |||||
QY 781 GCGGCTGTTGGAAACGTTCAGACCGCGGCTGTGGCTATGCAAGTACGCGGACCGGGCGGG 840
DB |||||
QY 781 GCGGCTGTTGGAAACGTTCAGACCGCGGCTGTGGCTATGCAAGTACGCGGACCGGGCGGG 840
DB |||||
QY 841 CCTACTCCCGGGTGTCTGCGGCGGAGAGGCTGCGCGCTCTCTCAGCGGAGACGG 900
DB |||||
QY 841 CCTACTCCCGGGTGTCTGCGGCGGAGAGGCTGCGCGCTCTCTCAGCGGAGACGG 900
DB |||||
QY 901 GTTTCGTGGAGAGACGACCCCGCGGGTGGAGCCCGGGGCTTCCCTGAAACCTTCCAGGC 960
DB |||||
QY 901 GTTTCGTGGAGAGACGACCCCGCGGGTGGAGCCCGGGGCTTCCCTGAAACCTTCCAGGC 960
DB |||||
QY 961 CACCGGCTTCCCGCCACGCTGCCACCGGACCTTTCGCGGGGCGCATTCAGAGCGCTG 1020
DB |||||
QY 961 CACCGGCTTCCCGCCACGCTGCCACCGGACCTTTCGCGGGGCGCATTCAGAGCGCTG 1020
DB |||||
QY 1021 CTGACCGCTCACAACGAGGGCCCTGGAGCGGCGCCACCGGCGCAGGCGCGCTTCGAGG 1080
DB |||||
QY 1021 CTGACCGCTCACAACGAGGGCCCTGGAGCGGCGCCACCGGCGCAGGCGCGCTTCGAGG 1080
DB |||||
QY 1081 GTGCGTGGACTGTGTGCGGCAACCCCAACGAGAGCAGTGGAGCGGAGATCC 1132
DB |||||
QY 1081 GTGCGTGGACTGTGTGCGGCAACCCCAACGAGAGCAGTGGAGCGGAGATCC 1132
DB |||||
```

```
RESULT 2
US-10-621-113-1
; Sequence 1, Application US/10621113
; Publication No. US20040091466A1
; GENERAL INFORMATION:
; APPLICANT: Lambeth, J. David
; APPLICANT: Cheng, Guangjie
; TITLE OF INVENTION: Regulatory Protein For Nox Enzymes
; FILE REFERENCE: 05501-0202 (43150-287577)
; CURRENT APPLICATION NUMBER: US/10/621.113
; CURRENT FILING DATE: 2003-07-16
; PRIOR APPLICATION NUMBER: US 60/405,647
; PRIOR FILING DATE: 2002-08-23
; PRIOR APPLICATION NUMBER: US 60/396,170
; PRIOR FILING DATE: 2002-07-16
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1129
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (5)..(1114)
; OTHER INFORMATION:
US-10-621-113-1
```

Query Match 98.6%; Score 1116; DB 17; Length 1129;
Best Local Similarity 99.7%; Pred. No. 3.1e-276;
Matches 1129; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

```
QY 1 AGCCATGGCAGGCCCCCGATACCCAGTTTCAGTGCAAGGGGCGAGCCCTGGTGCAATCAA 60
DB |||||
QY 1 AGCCATGGCAGGCCCCCGATACCCAGTTTCAGTGCAAGGGGCGAGCCCTGGTGCAATCAA 60
DB |||||
QY 61 GAGGCTCCAAAGCTTTGCTTCTCTGTGGCTGGTTCAGACGGCAGCGACACCTTCGTGG 120
DB |||||
QY 61 GAGGCTCCAAAGCTTTGCTTCTCTGTGGCTGGTTCAGACGGCAGCGACACCTTCGTGG 120
DB |||||
QY 121 CAGGAGTTGGGACGAATTCAGGACGCTCAAGAAGACCTCAAGGAGACCTTCCCGGTGGA 180
DB |||||
QY 121 CAGGAGTTGGGACGAATTCAGGACGCTCAAGAAGACCTCAAGGAGACCTTCCCGGTGGA 177
DB |||||
QY 181 GCGGGCTGTCTGCGGAGATCTGACCGGTTCTCCAAAGCTTCTCGATGCAACCACTGTT 240
DB |||||
QY 178 GCGGGCTGTCTGCGGAGATCTGACCGGTTCTCCAAAGCTTCTCGATGCAACCACTGTT 237
DB |||||
QY 241 GGGACGCTGGGGCCGACGAGCCCGCTGCGCGCTGCGAGCTGTTGGAAACCTATTTC 300
DB |||||
QY 238 GGGACGCTGGGGCCGACGAGCCCGCTGCGCGCTGCGAGCTGTTGGAAACCTATTTC 297
DB |||||
QY 301 TCGGAGGCTGTGCGGACTGCGAGCGCTGCGAGCGCGGACCGGACCGGATCACTGGCTTCTT 360
DB |||||
QY 298 TCGGAGGCTGTGCGGACTGCGAGCGCTGCGAGCGCGGACCGGACCGGATCACTGGCTTCTT 357
DB |||||
QY 361 GCGACCGCAACCCCTGGACCTGGAGCGCGCTGCGACCGGGGCGGGTGCCTGCC 420
DB |||||
QY 358 GCGACCGCAACCCCTGGACCTGGAGCGCGCTGCGACCGGGGCGGGTGCCTGCC 417
DB |||||
QY 421 CACCCAGAGGAGCAGCTTCTTCTGCGGCTGCGGGCGCGCTCTCCATCCACAGTCTGGA 480
DB |||||
QY 418 CACCCAGAGGAGCAGCTTCTTCTGCGGCTGCGGGCGCGCTCTCCATCCACAGTCTGGA 477
DB |||||
QY 481 GGCTCAGAGCTGCGCTGCTGCAAGCCCTTCTGTATCCAGGACACCGGGGATAGGCCCTTT 540
DB |||||
QY 478 GGCTCAGAGCTGCGCTGCTGCAAGCCCTTCTGTATCCAGGACACCGGGGATAGGCCCTTT 537
DB |||||
QY 541 TCAGGCGGAGGCCACAGGAGACCTTGGAGCTGCTGCTGGGCGACCCCTCAGGCTGGTGGCT 600
DB |||||
QY 538 TCAGGCGGAGGCCACAGGAGACCTTGGAGCTGCTGCTGGGCGACCCCTCAGGCTGGTGGCT 597
DB |||||
QY 601 GGTGAGAAACGAGACCGGCGAGACCGCTGTGTTTTCCAGCGCCCTACCTGGAGGAGGCGGC 660
```

```
Db 598 GGTGGAAAGAAAGACCGGAGACCGGCTGTGTTTCCAGCGCCCTACCTGGAGAGGCGGC 657
Qy 661 CCCGGGCAAGAGCGGAGGAGGAGCGCCGTCCTAGGAGAGAGCGGTCCCCAGTTCCTGTC 720
Db 658 CCCGGGCAAGAGCGGAGGAGGAGCGCCGTCCTAGGAGAGAGCGGTCCCCAGTTCCTGTC 717
Qy 721 TTCCCGCGCTACGAGAGAGCGGCGAGATGAGCTGCTCCGCGCGGCGGCGCGCT 780
Db 718 TTCCCGCGCTACGAGAGAGCGGCGAGATGAGCTGCTCCGCGCGGCGGCGCGCT 777
Qy 781 GCGCGTGTGGAAACGTCAGACCGCGCTGGTGGCTATGAGTACGGCCACCGGCGGG 840
Db 778 GCGCGTGTGGAAACGTCAGACCGCGCTGGTGGCTATGAGTACGGCCACCGGCGGG 837
Qy 841 CCTACTCCCCCGGTGTGCTGGCGCGGAGAGGCTGGGCGCTCTCTGAGCGGAGCGGG 900
Db 838 CCTACTCCCCCGGTGTGCTGGCGCGGAGAGGCTGGGCGCTCTCTGAGCGGAGCGGG 897
Qy 901 GTTCCGTGGAGAGAGACACCGCGGCGGTGAGGCGCGGGCTTCCCTGAACCTTCCAGGC 960
Db 898 GTTCCGTGGAGAGAGACACCGCGGCGGTGAGGCGCGGGCTTCCCTGAACCTTCCAGGC 957
Qy 961 CAGCGCCCTTCCCCACCGTGGCCACCGTGGCCACCGTGGCCACCGTGGCCACCGTGG 1020
Db 958 CAGCGCCCTTCCCCACCGTGGCCACCGTGGCCACCGTGGCCACCGTGGCCACCGTGG 1017
Qy 1021 CTGCACCGTCACACGAGGAGGCGCTGGAGCGGCGGCCACCGCGCGCGCGCGCTGGAG 1080
Db 1018 CTGCACCGTCACACGAGGAGGCGCTGGAGCGGCGGCCACCGCGCGCGCGCGCTGGAG 1077
Qy 1081 GTGCGTGGACTCTGTGCGGACCGCCACCGAGCGAGCACTGAGCGGAGGATCC 1132
Db 1078 GTGCGTGGACTCTGTGCGGACCGCCACCGAGCGAGCACTGAGCGGAGGATCC 1129
```

RESULT 3

```
US-10-621-113-5
; Sequence 5, Application US/10621113
; Publication No. US20040091466A1
; GENERAL INFORMATION:
; APPLICANT: Lambeth, J. David
; APPLICANT: Cheng, Guangjie
; TITLE OF INVENTION: Regulatory Protein For Nox Enzymes
; FILE REFERENCE: 05501-0202 (43150-287577)
; CURRENT APPLICATION NUMBER: US/10/621.113
; CURRENT FILING DATE: 2003-07-16
; PRIOR APPLICATION NUMBER: US 60/405,647
; PRIOR FILING DATE: 2002-08-23
; PRIOR APPLICATION NUMBER: US 60/396,170
; PRIOR FILING DATE: 2002-07-16
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 5
; LENGTH: 1147
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (5)..(1132)
; OTHER INFORMATION:
```

```
US-10-621-113-5
Query Match 97.8%; Score 1107; DB 17; Length 1147;
Best Local Similarity 98.7%; Pred. No. 6.3e-274;
Matches 1132; Conservative 0; Mismatches 0; Indels 15; Gaps 1;
Qy 1 AGCCATGGCAGGCGCCCGATACCCAGTTTCAGTGCAGGAGGCGAGCCCTGGTGCAGATCAA 60
Db 1 AGCCATGGCAGGCGCCCGATACCCAGTTTCAGTGCAGGAGGCGAGCCCTGGTGCAGATCAA 60
Qy 61 GAGGCTCCAAACGTTTGGCTTCTCTGTGGCTGGTTCAGAGCGGAGCGACACCTTCGTGGC 120
```

```
Db 61 GAGGCTCCAAACGTTTGGCTTCTCTGTGGCTGGTTCAGAGCGGAGCGACACCTTCGTGGC 120
Qy 121 CAGGAGTTGGGACGAATTCAGGCGAGCTCAAGAAAGACCCCTCAAGAGAGACCTTCCCGGTGGA 180
Db 121 CAGGAGTTGGGACGAATTCAGGCGAGCTCAAGAAAGACCCCTCAAGAGAGACCTTCCCGGTGGA 180
Qy 181 GGCGGCGCTGCTCGGAGATCTGACCGCGTTTCTCCAAAGCTTCTC----- 226
Db 181 GGCGGCGCTGCTCGGAGATCTGACCGCGTTTCTCCAAAGCTTCTC----- 240
Qy 227 -GATGCACACCTGTTTGGGAGCGGTGGGCGCAGAGCGCGGCGCTGGCGGCTTCGAGCT 285
Db 241 GGATGCACACCTGTTTGGGAGCGGTGGGCGCAGAGCGCGGCGCTGGCGGCTTCGAGCT 300
Qy 286 GTTGGAAAACCTATCTCGGAGGCTGTGGCGACTGCAGAGCGCTGGCAGGAGCGCGAC 345
Db 301 GTTGGAAAACCTATCTCGGAGGCTGTGGCGACTGCAGAGCGGTGGCAGGAGCGCGAC 360
Qy 346 GATCACTGGGTTCTTCGCAACCGCAACCTTGACCTGGAGCGCGCTGCCACCCGCGCAG 405
Db 361 GATCACTGGGTTCTTCGCAACCGCAACCTTGAGACTGGAGCGCGCTGCCACCCGCGCAG 420
Qy 406 CCGGTTGATCTGCTGCCACCGCAGAGGAGCGCTTCTTCGCGGCTTCGCGGCGCGCTCTC 465
Db 421 CCGGTTGATCTGCTGCCACCGCAGAGGAGCGCTTCTTCGCGGCTTCGCGGCGCGCTCTC 480
Qy 466 CATCCACAGTCTGGAGGCTCAGAGCGCTGCGCTGCGCTGAGCGGCTTCTGTACCCAGGACAC 525
Db 481 CATCCACAGTCTGGAGGCTCAGAGCGCTGCGCTGCGCTGAGCGGCTTCTGTACCCAGGACAC 540
Qy 526 GCGGATAGGCGCTTTCAGGCGCAGGCCCAGGAGAGCGCTTGGACGCTGCTGTGCGGCAACC 585
Db 541 GCGGATAGGCGCTTTCAGGCGCAGGCCCAGGAGAGCGCTTGGACGCTGCTGTGCGGCAACC 600
Qy 586 CTCAGGCTGGTGGCTGGAGAACGAAGACCGGAGAGCGCTTGGTTCAGCGGCGCTA 645
Db 601 CTCAGGCTGGTGGCTGGAGAACGAAGACCGGAGAGCGCTTGGTTCAGCGGCGCTA 660
Qy 646 CTTGAGAGGAGGCGGCGCCCGGCGCAAGCGCGGAGGAGGCGGCTTCTTAGGAGAGGAGCGG 705
Db 661 CTTGAGAGGAGGCGGCGCCCGGCGCAAGCGCGGAGGAGGCGGCTTCTTAGGAGAGGAGCGG 720
Qy 706 TCCCCAGTTCTGTGCTTCCCGCGCTTACGAGAGCAGCCCGCAGATGAGCTGTCCGTGCC 765
Db 721 TCCCCAGTTCTGTGCTTCCCGCGCTTACGAGAGCAGCCCGCAGATGAGCTGTCCGTGCC 780
Qy 766 CCGCGGGGCGCGCGCTGCTGTGGAACAGTGCAGCCGCGGCTGGTGGCTATGCAAGTTA 825
Db 781 CCGCGGGGCGCGCGCTGCTGTGGAACAGTGCAGCCGCGGCTGGTGGCTATGCAAGTTA 840
Qy 826 CCGCGACCGGGCGGCGCTACTCTCCCGCGGTGCTGCTGCGCGGAGGAGGCTGGGCGCTCT 885
Db 841 CCGCGACCGGGCGGCGCTACTCTCCCGCGGTGCTGCTGCGCGGAGGAGGCTGGGCGCTCT 900
Qy 886 CTTGAGCGGAGCGGCGTTCCGTGGAGAGACGACCCGCGGCGGTGAGGCGCCGCGGCTTCCC 945
Db 901 CTTGAGCGGAGCGGCGTTCCGTGGAGAGACGACCCCGCGGCGGTGAGGCGCCGCGGCTTCCC 960
Qy 946 TGAACCCCTCCAGGCGCACCGCGCTTCCCGCGCGCTGCGCGGAGGAGGCTTCCGCGGCGCG 1005
Db 961 TGAACCCCTCCAGGCGCACCGCGCTTCCCGCGCGCTGCGCGGAGGAGGCTTCCGCGGCGCG 1020
Qy 1006 CATCCAGAGCGCTGCTGCACCGTTCAGAGCGAGGCGCTTGGAGCGGCGCCCGCAGCGCGCA 1065
Db 1021 CATCCAGAGCGCTGCTGCACCGTTCAGAGCGAGGCGCTTGGAGCGGCGCCCGCAGCGCGCA 1080
Qy 1066 GGGCGCGCTTCCAGGAGTGGCTGGACTCTGTGCGGCAACCCACGACGAGGAGTGGAGCGG 1125
Db 1081 GGGCGCGCTTCCAGGAGTGGCTGGACTCTGTGCGGCAACCCACGACGAGGAGTGGAGCGG 1140
Qy 1126 AGGATCC 1132
Db 1141 AGGATCC 1147
```

RESULT 4

US-10-621-113-7
; Sequence 7, Application US/10621113
; Publication No. US20040091466A1
; GENERAL INFORMATION:
; APPLICANT: Lambeth, J. David
; APPLICANT: Cheng, Guangjie
; TITLE OF INVENTION: Regulatory Protein For Nox Enzymes
; FILE REFERENCE: 05501-0202 (43150-287577)
; CURRENT APPLICATION NUMBER: US/10/621,113
; CURRENT FILING DATE: 2003-07-16
; PRIOR APPLICATION NUMBER: US 60/405,647
; PRIOR FILING DATE: 2002-08-23
; PRIOR APPLICATION NUMBER: US 60/396,170
; PRIOR FILING DATE: 2002-07-16
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 1144
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (5)..(1129)
; OTHER INFORMATION:
US-10-621-113-7

Query Match 96.4%; Score 1091; DB 17; Length 1144;
Best Local Similarity 98.4%; Pred. No. 7.9e-270;
Matches 1129; Conservative 0; Mismatches 0; Indels 18; Gaps 2;

QY 1 AGCCATGGCAGGCCCCCGATACCAAGTTTCAGTGAAGGGGAGCCCTGGTGCAATCAA 60
Db 1 AGCCATGGCAGGCCCCCGATACCAAGTTTCAGTGAAGGGGAGCCCTGGTGCAATCAA 60

QY 61 GAGGCTCCAAAGTTTGGCTTCTCTGTGGCTGGTTCAGACGGCAGACACCTTGGTGG 120
Db 61 GAGGCTCCAAAGTTTGGCTTCTCTGTGGCTGGTTCAGACGGCAGACACCTTGGTGG 120

QY 121 CAGGATGGGACGAATTCAGGACGCTCAAGAGACCTTCAAGAGACCTTCCCGTGGGA 180
Db 121 CAGGATGGGACGAATTCAGGACGCTCAAGAGACCTTCAAGAGACCTTCCCGTGGGA 177

QY 181 GCGGGCTGTCTGCGGAGATCTGACCGGTTCCTCCAAAGCTTCTCGGTTCAGGCCAGCCT 226
Db 178 GCGGGCTGTCTGCGGAGATCTGACCGGTTCCTCCAAAGCTTCTCGGTTCAGGCCAGCCT 237

QY 227 GATCACCACCTGTTGGGACGGTGGGGGACGAGCCCGGCTGGCGCGCTGCAGCT 285
Db 238 GATCACCACCTGTTGGGACGGTGGGGGACGAGCCCGGCTGGCGCGCTGCAGCT 297

QY 286 GTTGAACACCTATTCTCGGAGCTGTGCGGAGCTGCGGAGCGCTGCGACGAGCGCCGAC 345
Db 298 GTTGAACACCTATTCTCGGAGCTGTGCGGAGCTGCGGAGCGCTGCGACGAGCGCCGAC 357

QY 346 GATCACTGGCTTCTTGCACCGCAACCCCTGGACCTTGAGCCCGGCTGCCACCCGGGAG 405
Db 358 GATCACTGGCTTCTTGCACCGCAACCCCTGGACCTTGAGCCCGGCTGCCACCCGGGAG 417

QY 406 CCGGGTATCTGCGCCACCCAGAGGAGCGCTCTTCTCGCGTGGCGGCGGCTCTC 465
Db 418 CCGGGTATCTGCGCCACCCAGAGGAGCGCTCTTCTCGCGTGGCGGCGGCTCTC 477

QY 466 CATCACAGCTTGGAGGCTCAGAGCTGCGCTGCTGAGCCCTTCTGTACCCAGGACAC 525
Db 478 CATCACAGCTTGGAGGCTCAGAGCTGCGCTGCTGAGCCCTTCTGTACCCAGGACAC 537

QY 526 GCGGATAGGCTTTTCAGGCGCAGGCCAGAGAGCTTGAAGCTGTGCTGCGGCAACC 585
Db 538 GCGGATAGGCTTTTCAGGCGCAGGCCAGAGAGCTTGAAGCTGTGCTGCGGCAACC 597

QY 586 CTCAGGCTGGTGGCTGGAGAAAGAAAGACGGGAGAGCCGCTGTTTCCAGCGCCCTTA 645
Db 598 CTCAGGCTGGTGGCTGGAGAAAGAAAGACGGGAGAGCCGCTGTTTCCAGCGCCCTTA 657

QY 646 CTTGAGAGAGCGGCGCCCGGGCCAAAGCGGGAGAGAGCCCGTCCCTTAGGAGAGAGCGG 705
Db 658 CTTGAGAGAGCGGCGCCCGGGCCAAAGCGGGAGAGAGCCCGTCCCTTAGGAGAGAGCGG 717

QY 706 TCCCAGTTCTGTGCTTCCCGGCTTACGAGAGAGCGCGGAGATGAGCTGTCTCGTGCC 765
Db 718 TCCCAGTTCTGTGCTTCCCGGCTTACGAGAGAGCGCGGAGATGAGCTGTCTCGTGCC 777

QY 766 CGCGGGCGCGCGCTGCGCTGTTGAAACGTGACACCGCGCTGCTGGCTATGACAGGTA 825
Db 778 CGCGGGCGCGCGCTGCGCTGTTGAAACGTGACACCGCGCTGCTGGCTATGACAGGTA 837

QY 826 CGCGAGCCGGCGGCGCTTACTCCCGCGGCTGTGCTGCGGCGGAGAGGCTGGGCGCTCT 885
Db 838 CGCGAGCCGGCGGCGCTTACTCCCGCGGCTGTGCTGCGGCGGAGAGGCTGGGCGCTCT 897

QY 886 CTTGAGCGGAGAGCGGCTTCCGTTGAGAGAGAGACCCCGGCGGCTGAGGCGCGGCTTCCC 945
Db 898 CTTGAGCGGAGAGCGGCTTCCGTTGAGAGAGAGACCCCGGCGGCTGAGGCGCGGCTTCCC 957

QY 946 TGAACCTCCAGAGCCAGCGCCCTCCCGCCACCGTGCACCCGACCTTGCCTGGGCGC 1005
Db 958 TGAACCTCCAGAGCCAGCGCCCTCCCGCCACCGTGCACCCGACCTTGCCTGGGCGC 1017

QY 1006 CATCCAGAGCGCTGCTGCACCGCTCACAAGAGGCGCTTGGAGCGGCGCCCAAGCGCGCA 1065
Db 1018 CATCCAGAGCGCTGCTGCACCGCTCACAAGAGGCGCTTGGAGCGGCGCCCAAGCGCGCA 1077

QY 1066 GGGCGCGCTCAGAGGCTGCTGAGCTCTGTGCGGACCCCGACGAGAGAGTGAAGCGG 1125
Db 1078 GGGCGCGCTCAGAGGCTGCTGAGCTCTGTGCGGACCCCGACGAGAGAGTGAAGCGG 1137

QY 1126 AGGATCC 1132
Db 1138 AGGATCC 1144

RESULT 5

US-10-108-260A-1465/c
; Sequence 1465, Application US/10108260A
; Publication No. US20040005560A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20040005560A1el full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/10/108,260A
; CURRENT FILING DATE: 2002-03-27
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1465
; LENGTH: 2922
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-108-260A-1465

Query Match 30.2%; Score 342.4; DB 17; Length 2922;
Best Local Similarity 84.9%; Pred. No. 3.9e-78;
Matches 428; Conservative 0; Mismatches 1; Indels 75; Gaps 1;

QY 704 GGTCCCAGATTTCTGTGCTTCCCGGCTTACAGAGAGAGCCCGGAGATGAGCTGTCCGTG 763
Db 679 GGTCCCAGATTTCTGTGCTTCCCGGCTTACAGAGAGAGCCCGGAGATGAGCTGTCCGTG 620

QY 764 CCCGCGGGGCGCGCTGCGCTTGGAAACGTGACACCGCGCTGGTGCTAT----- 818
Db 619 CCCGCGGGGCGCGCTGCGCTTGGAAACGTGACACCGCGCTGGTGCTATGACAGG 560

QY 819 ----- 818

Db 559 TACCGGGAGCGGGCGTGGGAGGGCAGGCTAGCCGAGGGGGGGGGCCCTAACACAC 500
Qy 819 -----GCAGGTACGGCGACCGGGCGGGCCTACTCCCGGGGTGCTGCTGCGGCCG 868
Db 499 CCGCGCCCTCGCAGGTACGGGACCGGGCGGGCTACTCCCGGGGTGCTGCTGCGGCCG 440
Qy 869 GAAGGGTGGGGCTTCTTGAAGCGGACCGGGTTCGTTGGAGGAGACGACCGGGCGGGT 928
Db 439 GAAGGGTGGGGCTTCTTGAAGCGGACCGGGTTCGTTGGAGGAGACGACCGGGCGGGT 380
Qy 929 GAGSCCGGGGCTTCCCTGTAACCTCCAGGCGACCGCCCTCCCGCCACGGTGCCAC 988
Db 379 GAGSCCGGGGCTTCCCTGTAACCTCCAGGCGACCGCCCTCCCGCCACGGTGCCAC 320
Qy 989 CGACCTTCGCGGGCGGCATCCAGAGCGGCTGTGTGACCGCTCACACAGGGCGCTTGAG 1048
Db 319 CGACCTTCGCGGGCGGCATCCAGAGCGGCTGTGTGACCGCTCACACAGGGCGCTTGAG 260
Qy 1049 CGGGGCCACCGGCCACGGCGCGCTCGAGGGTGGCTGACTTGTGCGGCACCGCCACG 1108
Db 259 CGGGGCCACCGGCCACGGCGCGCTCGAGGGTGGCTGACTTGTGCGGCACCGCCACG 200
Qy 1109 ACCGAGCAGTGAGCGCGAGGATCC 1132
Db 199 ACCGAGCAGTGAGCGCGAGGATCC 176

RESULT 6
US-10-437-963-24598
; Sequence 24598, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; CURRENT APPLICATION NUMBER: US/10/437,963
; PRIOR FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 24598
; LENGTH: 1180
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(1180)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_29567C.1
US-10-437-963-24598

Query Match 5.3%; Score 60; DB 18; Length 1180;
Best Local Similarity 48.7%; Pred. No. 8.9e-06;
Matches 170; Conservative 0; Mismatches 178; Indels 1; Gaps 1;
Qy 759 CCGTGCCCGGGGGCGCGGTGCGGTGTGAAACGTGAGACCGCGGCTGTGCTAT 818
Db 308 CCGGGGGCGCGGGNCCCGCGCCCGCGGGGCGGGGCGCCCGCCCGCCCGCGCG 367
Qy 819 GCAGGTACCGCGACCGGGCGGGCTTACTCCCGCGGTGCTGCTGCGCGCGAAGGGCTGG 878
Db 368 CGSGCGCNGGCGCCNCCCGCGCC-GCGCGCGCGGGCGCCCGCGCGGGCGCGCGCG 426
Qy 879 GCGCTCTCTGACGGGACGGGGTTTCGTGAGGAGACGACCGCGGGGTGAGCGCCGGG 938
Db 427 GCCCGCGCGGGGGCGCGCGCGCCCGCGGGCGGGCGCGCGCGCGCGCGCGCGCG 486

Qy 939 GCTTCTTGAACCTTCCAGGCGACCGGCCCTTCCCGCCACCGTGCACCGACCTTGC 998
Db 487 CCGCGGGCG 546
Qy 999 CGGGCGGCATTCAGAGCGGCTGTGTGACCGTCAACGAGGGCCCTGAGAGCGCGCCAC 1058
Db 547 CG 606
Qy 1059 GGGCGCAGGCGCGCGCTCGAGGGTGGTGAAGTCTGTGCGCGCACCGCCAC 1107
Db 607 CCNNGCG 655

RESULT 7
US-10-156-761-7099
; Sequence 7099, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 7099
; LENGTH: 2076
; TYPE: DNA
; ORGANISM: Streptomyces avermitilis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2076)
US-10-156-761-7099

Query Match 5.1%; Score 58.2; DB 15; Length 2076;
Best Local Similarity 44.0%; Pred. No. 2.5e-05;
Matches 380; Conservative 0; Mismatches 478; Indels 5; Gaps 3;
Qy 263 CGCGGCTGCGCGCGCTGCGAGCTGTTGGAAACCTATTCTCGGAGGCTGTGCGGACTGCA 322
Db 424 CGCGGCATCATGCACCGGGACGTGAAGCGCGGAACATCTGCTGCGCGGACCGCACG 483
Qy 323 GAGCGGTGGCAGGAGCCGCGAGCATCACTGGGTTCTTGCACCGCAACCCCTGGACCTG 382
Db 484 GGGCGCGGTACGCGCGCGTGTCTCACCGACTACGGCATCTCCGTGCAACCGGACGCG 543
Qy 383 GAGCCCGCGCTGCCACCGCGGAGCGGGTGATCTCCGCCACCGAGGAGGAGCGCTCTT 442
Db 544 GGGAGACCGGTCACACCTCAGTCGCGCGCTGTGTCGGACGCGCCCGGCTATCTGGGCGG 603
Qy 443 TCTCGCGCTGCGGGCGCGCTCTCCATCCACAGTGTGAGGCTCAGAGCCTGCGCTCCTG 502
Db 604 GAGCGGGACCGCGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 663
Qy 503 CAGCCCTTGTGTACCGAGGACAGCGGGGATAGGCTTTTTCAGGCGGAGCGCCAGGAGAGC 562
Db 664 CTGTACTTCGGGGTTCGAGGGGTGCGGGCGCTTCGAGCGGGGACACGCTCTCGCCGAGGC 723
Qy 563 CTGGAGCTGCTGTCGCGCACCCCTCAGGCTGCTGCTGCTGAGAGAACGAGACCGGCGAG 622
Db 724 ACCGCGGTGTGTGAGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGGTG 783
Qy 623 ACCGCGGTGTTTCCAGCGCCCTACCTGAGGAGGCGCGCGCGCGCGCGCGCGCGCGG 682

RESULT 10
US-10-357-930-23327
; Sequence 23327, Application US/10357930
; Publication No. US20040259086A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Endege, Wilbon
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; TITLE OF INVENTION: HUMAN PROSTATE CANCER
; FILE REFERENCE: MRI-007BCN
; CURRENT APPLICATION NUMBER: US/10/357,930
; CURRENT FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: 09/785,276

Query Match	5.0%;	Score 56.8;	DB 18;	Length 2781;
Best Local Similarity	47.8%;	Pred. No. 5.5e-05;		
Matches 198;	Conservative 0;	Mismatches 212;	Indels 4;	Gaps 1
Qy	205	CCGGGTCTCCAAAGCTTCTCGATGCACCACTGTTGGAGCGCTGGGGCGCACAGAGCGG	264	
Db	17	CGGGTTCTCCGGGGCTACGCGGGGCCCTGGGCCAGGTTGGCTGCCAGGCTCGGCGG	76	
Qy	265	CGGCTGGCGCGCTGCAGCTGTTGGAAACCTATTCTCGAGGGTGTGGCGCATGTCAGA	324	
Db	77	GAGCGTGGAGCCCGCGCGCTGCCCCAGGAC---CGCGCCCGCGCCTTTGTCCGCGCCG	132	
Qy	325	GCGGTGGCAGGAGCCCGACGATCACTGGCTTCTTCGCACCGCAACCCCTGGACCTTGA	384	
Db	133	GCCACCGCCCGTCCCCCGCCGCCATGAGAGCGCCGCGCGCTGCGCGCGGGTCCCGCT	192	
Qy	385	GCCCGCGCTGCCACCCGGCAGCGGGTGATCTGCCCCACCCAGAGGAGCAGCCTCTTTC	444	
Db	193	TCCGCTGCTGCTGCTCGGCGGCCCTTGCGTCTGCGCGCCGGAGTGGACCGCGATGCTCT	252	
Qy	445	TCGCGCTGGGGCGCCCTCTCATTCACAGTCTGGAGGCTCAGAGCCTGCGCTGCTGCA	504	
Db	253	CCTGGAGGCTGTGTGTGGCAGCGACACCCGATGSCCACTCATCAGAAGGACTGCTCGCT	312	
Qy	505	GCCCTTCTGTACCCAGGACACGGGATAGCGCTTTTTCAGCGCAGGCCCCAGGAGAGCT	564	
Db	313	GCCATATGCTACGGAAATCCAAAGNATGCAGATGGTTCAGAGACATGCTGCCACAGCA	372	
Qy	565	GGAGCTGCTGCTCGGCACCCCTCAGGCTGGTGGCTGGTGGAGAACGAAGACCG	618	
Db	373	GCTGGAGAGCTGCATGTGCTGCCACGGGCATCAGCTGTGCCCAACGAGCAGGACCG	426	

```

RESULT 11
US-10-357-930-29208
; Sequence 29208, Application US/10357930
; Publication No. US20040259086A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Endege, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; TITLE OF INVENTION: HUMAN PROSTATE CANCER
; FILE REFERENCE: MRI-007BCN
; CURRENT APPLICATION NUMBER: US/10/357,930
; CURRENT FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: 09/785,276

```

```
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq For Windows Version 4.0
; SEQ ID NO 29208
; LENGTH: 2781
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1, 2, 3, 4, 2779, 2780, 2781
; OTHER INFORMATION: n = A,T,C or G
US-10-930-29208

Query Match          5.0%; Score 56.8; DB 18; Length 2781;
Best Local Similarity 47.8%; Pred. No. 5.5e-05;
Matches 198; Conservative 0; Mismatches 212; Indels 4; Gaps 1;

QY 205 CGCGCTTCTCCCAAGCTTCTCGATGCACCACTGTGTGGACGCGGTGGGCGCACGAGCG 264
DB 17 CGGGTTCTCCCGGGCTACGGCGGCTGTGCCAGAGTTGGCTGCCGAGGCTCGGCG 76

QY 265 CGCGCTGGCGGCGCTGCACTGTGTGGAAACCTATTCTCGAGGCTGTGGCACTGCAGA 324
DB 77 GAGCGTGGAGCGCGCGCTGTGCCAGGAC----CGCGCGCGCTTGTTCGCGCGCC 132

QY 325 GCGGTGGCAAGGACCGACGATCACTGGCTTCTTGGACCGCAACCTCGGACCTGGA 384
DB 133 GCCACCGCGCTGCGCGCGCGCCATGAGCGCGCGCGCTGCGCGCGCGGTCCCGCT 192

QY 385 GCCCGCTGCGCACCGCGCGGTGATCTGCGCCACCGCGCGCGCGCGCGCGCGCTTTC 444
DB 193 TCCGCTGTCTGCTCGGGGCGCTTGGCTGTGGCGCGCGCGCGCGCGCGCGCGATGCT 252

QY 445 TCGCGCTCGGCGCGCTTCTCCATCCACAGTCTGAGGCTCAGAGCTGCGCTGCTGCA 504
DB 253 CTGGAGGCGCTGTGCGGACGAGACCGGATGCCACTCATCAGAAGGACTGCTCGCT 312

QY 505 GCCCTTCTGTACCCAGGACCGCGGATAGGCTTTTTCAGGCGGAGCGCGCGGAGGCT 564
DB 313 GCCATATGTACGGAATCCAAAGATGCAAGATGTCAGGATGTCAGGAGCAGTGTGCA 372

QY 565 GAGCTGTCTGTCGCGCGCGCGCTCAGGCTGCGGTGTGGTGGAGAACGAGACCG 618
DB 373 GCTGAGGAGCTGCATCTGTGCCACGGGATCAGCTTGGCCACGAGGAGACCG 426

RESULT 12
US-10-425-115-94529/c
; Sequence 94529, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326

; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq For Windows Version 4.0
; SEQ ID NO 29208
; LENGTH: 2781
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1, 2, 3, 4, 2779, 2780, 2781
; OTHER INFORMATION: n = A,T,C or G
US-10-930-29208

Query Match          5.0%; Score 56.8; DB 18; Length 2781;
Best Local Similarity 47.8%; Pred. No. 5.5e-05;
Matches 198; Conservative 0; Mismatches 212; Indels 4; Gaps 1;

QY 205 CGCGCTTCTCCCAAGCTTCTCGATGCACCACTGTGTGGACGCGGTGGGCGCACGAGCG 264
DB 17 CGGGTTCTCCCGGGCTACGGCGGCTGTGCCAGAGTTGGCTGCCGAGGCTCGGCG 76

QY 265 CGCGCTGGCGGCGCTGCACTGTGTGGAAACCTATTCTCGAGGCTGTGGCACTGCAGA 324
DB 77 GAGCGTGGAGCGCGCGCTGTGCCAGGAC----CGCGCGCGCTTGTTCGCGCGCC 132

QY 325 GCGGTGGCAAGGACCGACGATCACTGGCTTCTTGGACCGCAACCTCGGACCTGGA 384
DB 133 GCCACCGCGCTGCGCGCGCGCCATGAGCGCGCGCGCTGCGCGCGCGGTCCCGCT 192

QY 385 GCCCGCTGCGCACCGCGCGGTGATCTGCGCCACCGCGCGCGCGCGCGCGCGCTTTC 444
DB 193 TCCGCTGTCTGCTCGGGGCGCTTGGCTGTGGCGCGCGCGCGCGCGCGCGCGATGCT 252

QY 445 TCGCGCTCGGCGCGCTTCTCCATCCACAGTCTGAGGCTCAGAGCTGCGCTGCTGCA 504
DB 253 CTGGAGGCGCTGTGCGGACGAGACCGGATGCCACTCATCAGAAGGACTGCTCGCT 312

QY 505 GCCCTTCTGTACCCAGGACCGCGGATAGGCTTTTTCAGGCGGAGCGCGCGGAGGCT 564
DB 313 GCCATATGTACGGAATCCAAAGATGCAAGATGTCAGGATGTCAGGAGCAGTGTGCA 372

QY 565 GAGCTGTCTGTCGCGCGCGCGCTCAGGCTGCGGTGTGGTGGAGAACGAGACCG 618
DB 373 GCTGAGGAGCTGCATCTGTGCCACGGGATCAGCTTGGCCACGAGGAGACCG 426

RESULT 12
US-10-425-115-94529/c
; Sequence 94529, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
```

```
; SEQ ID NO 94529
; LENGTH: 1521
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_186203C.1
US-10-425-115-94529

Query Match          5.0%; Score 56.4; DB 18; Length 1521;
Best Local Similarity 46.4%; Pred. No. 7.3e-05;
Matches 218; Conservative 0; Mismatches 251; Indels 1; Gaps 1;

QY 646 CCTGGAGGAGCGCGCCCGCGGCCCAAGCGCGGAGGAGGCGCGCTCCCTTAGGAGCAGCG 705
DB 1004 CCAGACGTAGCAGTCTGCGGCGCGCTGAGCGCGGCTGTAGTAGTCTTCGCGCGC 945

QY 706 TCCCCAGTTCTGTCTTCCCGCGCTACGAGAGCAGCGCGCAGATGATGAGTGTCCGTGCC 765
DB 944 GGCCAGGCTCGCGGCTGTCTCCCAACGAGGCGCGCGTAGAGCAGCGCTCGGCGT 885

QY 766 CGCGGGGCGCGCTGCGCGGTGTGGAAACGTAGACCGCGGCTGTGGCTATGCAGTA 825
DB 884 CGCGGGGCTGCGGAGGATGCGCGCTCGCAGTACTCTGCGCGCTTTCGCGCGTCCCG 825

QY 826 CGCGAGCGCGCGCGCTTACTCCCGCGGTGCTGTCGCGCGGAAAGGCTGGGCGCTCT 885
DB 824 CCACCTCTTGAGAACCTGCGGTAGTTGACGAGGAGCAGGAGTTGCGCGGTCCGCT 765

QY 886 CCTGAGCGGAGCGGGGTTCCGTGAGGAGACGACCCCGCGGTGAGGCGCGGGGCTTCCC 945
DB 764 CGATCATCTGCGCGTAGTGGCGCTCCGTGGCGCGCGGTTGCTGCCACCGTCCGCTCCC 705

QY 946 TGAACCTCCAGGCGACCGCGCTTCCCGCCACCGTGGCCACCGACCTTCCCGGGGCG 1005
DB 704 CGACCCACCGCTTCCCGCGCTGCTGCGCGACGCGCTTCCCGCGCGCGCTGCGCGCG 645

QY 1006 CATCCAGAGCGCTGTCACCGTCACAGCAGGCGCTGAGGCGCGCGCCACGCGCGCA 1065
DB 644 CGCCTGTCTCAGAACGCGAGCGCGCGGCGCGGACACGTCCAGCGCGCTGTTCTGCA 585

QY 1066 G-GGCGCGCTTCGAGGCTGCGTGTGAGTCTGTGTCGCGCACCCCGACGAG 1114
DB 584 GCAGCGAGCGCGCGGACGCGCGCGTCCACGCGCGCGCGCGCGCGCGCGCGCG 535

RESULT 13
US-10-278-698-276
; Sequence 276, Application US/10278698
; Publication No. US20050037344A1
; GENERAL INFORMATION:
; APPLICANT: PathoArray GmbH
; APPLICANT: Stuhlmüller, Bruno
; APPLICANT: Haupt, Thomas
; TITLE OF INVENTION: Nucleic Acid Array
; FILE REFERENCE: 030027US
; CURRENT APPLICATION NUMBER: US/10/278,698
; CURRENT FILING DATE: 2002-10-23
; NUMBER OF SEQ ID NOS: 1050
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 276
; LENGTH: 2947
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-278-698-276

Query Match          4.9%; Score 55.6; DB 19; Length 2947;
Best Local Similarity 49.3%; Pred. No. 0.00011;
Matches 145; Conservative 0; Mismatches 149; Indels 0; Gaps 0;

QY 325 GCGGCTGGCAGCGAGCGCGAGTCACTGGCTTCTTGCACCGCAACCCCTGGACCTGGA 384
DB 123 GCCACCGCGCGCTGCGCGCGCGCCCATGGAGCGCGCGCTGCGCGCGCGCTCCCGCT 182
```

QY 385 GCCCGCTGTCACCCCGGAGCGGGTGATCTCTGCCACCCACAGAGAGAGCGCTCTTTC 444
 Db 183 TCCGCTGTGCTGTGCTGCGCGGCTTGGCGCTGCTGCGCGGAGTGCAGCGGATGTCCT 242
 QY 445 TCGGCTGCGGGCGGCTCTCCATCCACAGTCTGAGGCTCAGAGCCTGCGCTGCCGTGCA 504
 Db 243 CTTGAGGCTGTGTCGGACGACACCGGATGGCCACTCATCAGAGGACTGCTCGCT 302
 QY 505 GCCCTTCTGTACCCAGACACGCGGGATAGGCCCTTTTCAGGCGCAGGCCAGGAGCCT 564
 Db 303 GCCATATGCTACGGAATCCAAAGATGAGGATGGTGCAGGAGCAGTGTGCCACAGCCA 362
 QY 565 GGAGCTGCTGCTGGGACACCCCTCAGGCTGGTGGTGGTGGAGAACGAGACCG 618
 Db 363 GCTGGAGGAGCTGCATGTGCCACGGGCTATCAGCTTGGCCAAACGAGCAGGACCG 416

RESULT 14

US-10-278-698-790
 ; Sequence 790, Application US/10278698
 ; Publication No. US20050037344A1
 ; GENERAL INFORMATION:
 ; APPLICANT: PathoArray GmbH
 ; APPLICANT: Stuhlmüller, Bruno
 ; APPLICANT: Haupt, Thomas
 ; TITLE OF INVENTION: Nucleic Acid Array
 ; FILE REFERENCE: 030027US
 ; CURRENT APPLICATION NUMBER: US/10/278,698
 ; CURRENT FILING DATE: 2002-10-23
 ; NUMBER OF SEQ ID NOS: 1050
 ; SOFTWARE: Patentin version 3.2
 ; SEQ ID NO 790
 ; LENGTH: 2947
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-10-278-698-790

Query Match 4.9%; Score 55.6; DB 19; Length 2947;
 Best Local Similarity 49.3%; Pred. No. 0.00011;
 Matches 145; Conservative 0; Mismatches 149; Indels 0; Gaps 0;
 QY 325 GCGGTGGACAGGAGCGGAGTCACTGCTCTTGGACCGAACCCCTGACCTGGA 384
 Db 123 GCCACCGCGCTGCTGCGCGCGGCTTGGAGCGCGCGCGCTGCGCGGCTCCCGCT 182
 QY 385 GCGCGCTGCCACCGGAGCGGGTGATCTCTGCCACCCACAGAGGAGCAGCTCTTTC 444
 Db 183 TCCGCTGTGCTGCTGCGCGGCTTGGCTGCTGCGCGGCGGAGTGCAGCGGATGTCCT 242
 QY 445 TCGGCTGCGGGCGGCTCTTCCATCCACAGTCTGAGGCTCAGAGCCTGCGCTGCCGTGCA 504
 Db 243 CTTGAGGCTGTGTCGGACGACACCGGATGGCCACTCATCAGAGGACTGCTCGCT 302
 QY 505 GCCCTTCTGTACCCAGACACGCGGGATAGGCCCTTTTCAGGCGCAGGCCAGGAGCCT 564
 Db 303 GCCATATGCTACGGAATCCAAAGATGAGGATGGTGCAGGAGCAGTGTGCCACAGCCA 362
 QY 565 GGAGCTGCTGCTGGGACACCCCTCAGGCTGGTGGTGGTGGAGAACGAGACCG 618
 Db 363 GCTGGAGGAGCTGCATGTGCCACGGGCTATCAGCTTGGCCAAACGAGCAGGACCG 416

RESULT 15

US-10-425-114-31082/c
 ; Sequence 31082, Application US/10425114
 ; Publication No. US2004003488A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Liu, Jingdong
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Screen, Steven E.
 ; APPLICANT: Tabaska, Jack E.
 ; APPLICANT: Cao, Yongwei

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 ; FILE REFERENCE: 38-21(53313)B
 ; CURRENT APPLICATION NUMBER: US/10/425,114
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 73128
 ; SEQ ID NO 31082
 ; LENGTH: 1240
 ; TYPE: DNA
 ; ORGANISM: Zea mays
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: UC-2MFLB73125G09_FLI
 ; US-10-425-114-31082

Query Match 4.8%; Score 54.8; DB 17; Length 1240;
 Best Local Similarity 46.2%; Pred. No. 0.00019;
 Matches 217; Conservative 0; Mismatches 252; Indels 1; Gaps 1;
 QY 646 CTTGAGAGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 705
 Db 751 CCAGGACGTAGCAGTCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 692
 QY 706 TCCCGAGTTCTGTGCTTCCCGGCGCTACGAGACGACGCGGCGGCGGCGGCGGCGG 765
 Db 691 GGCACGCTGCGGCGTGGTCTCCAAACGAGGCGGCGGCGGCGGCGGCGGCGGCGG 632
 QY 766 CGCGGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 825
 Db 631 CGCGGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 572
 QY 826 CGCGGACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 885
 Db 571 CCACCTCTTGAGGAACCTGGCGTAGTTGACACGAGGAGCAGCGAGTTGCCCGGCT 512
 QY 886 CTTGAGCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 945
 Db 511 CGATCATCTGCGGCGTGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 452
 QY 946 TGAACCTCTCCAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1005
 Db 451 CGATCCACCTTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 392
 QY 1006 CATTCAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1065
 Db 391 CGGCTGTCTCCAGAACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 332
 QY 1066 G-GGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1114
 Db 331 GCAGCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 282

Search completed: May 31, 2005, 01:45:34
 Job time : 1024 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 30, 2005, 18:43:08 ; Search time 320 Seconds
(without alignments)
5788.329 Million cell updates/sec

Title: US-10-621-113-3
Perfect score: 1132
Sequence: 1 agccatggcagggccccgat.....agcagtgcgcgaggatccc 1132

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA: *
1: /cgn2_6/ptodata/1/ina/5A COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PCTUS COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfileseq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	50	4.4	971	3	US-08-630-915A-197
2	50	4.4	971	4	US-09-879-957-197
3	50	4.4	1272	4	US-09-949-016-3684
C 4	50	4.4	8438	1	US-07-945-283-1
5	50	4.4	42250	4	US-09-949-016-15426
C 6	49.4	4.4	152393	4	US-09-949-016-14514
C 7	49.4	4.4	152393	4	US-09-949-016-14515
C 8	49.4	4.4	156894	4	US-09-949-016-12765
C 9	49.4	4.4	156894	4	US-09-949-016-12766
C 10	49.4	4.4	156895	4	US-09-949-016-16957
C 11	49.4	4.4	156895	4	US-09-949-016-16958
C 12	49.4	4.4	156895	4	US-09-949-016-16959
13	49.2	4.3	801	2	US-08-770-379-16
14	49.2	4.3	801	3	US-08-757-669A-16
15	49.2	4.3	801	3	US-09-230-371A-16
C 16	49.2	4.3	3066	4	US-10-237-551-152
C 17	49.2	4.3	154746	4	US-09-827-688-8
C 18	49.2	4.3	154746	4	US-09-827-688-8
19	48.4	4.3	1457	3	US-09-444-053-3
C 20	47.8	4.2	19152	4	US-09-949-016-12110
C 21	47.8	4.2	19153	4	US-09-949-016-15795
C 22	47.8	4.2	90618	4	US-09-949-016-15964
C 23	47.6	4.2	801	3	US-09-298-568-3
C 24	47.6	4.2	801	3	US-09-894-273-3
C 25	47.4	4.2	76985	4	US-09-949-016-12416
C 26	47.4	4.2	76986	4	US-09-949-016-13120
27	46.8	4.1	1023	4	US-09-902-540-8340

28	46.8	4.1	2561	4	US-09-616-289-48	Sequence 48, Appl
C 29	46.8	4.1	8578	4	US-09-902-540-871	Sequence 871, App
30	46.6	4.1	1059	4	US-09-902-540-5064	Sequence 5064, Ap
C 31	46.6	4.1	3462	4	US-09-949-016-2642	Sequence 2642, Ap
C 32	46.6	4.1	3641	4	US-09-949-016-3877	Sequence 3877, Ap
C 33	46.6	4.1	7460	4	US-09-949-016-12375	Sequence 12375, A
C 34	46.6	4.1	7462	4	US-09-949-016-14384	Sequence 14384, A
C 35	46.6	4.1	7700	4	US-09-949-016-15619	Sequence 15619, A
C 36	46.6	4.1	28194	4	US-09-902-540-1250	Sequence 1250, Ap
C 37	46.6	4.1	152331	3	US-09-128-155-16	Sequence 16, Appl
C 38	46.4	4.1	1722	4	US-09-902-540-9668	Sequence 9668, Ap
39	46.4	4.1	14555	4	US-09-902-540-1096	Sequence 1096, Ap
40	46.4	4.1	71574	4	US-09-949-016-15580	Sequence 15580, A
41	46	4.1	9369	4	US-10-237-551-190	Sequence 190, App
42	46	4.1	9369	4	US-10-237-551-247	Sequence 247, App
43	45.8	4.0	1221	4	US-09-252-991A-8921	Sequence 8921, Ap
44	45.8	4.0	1545	4	US-09-252-991A-8710	Sequence 8710, Ap
45	45.8	4.0	2199	4	US-09-373-272-16	Sequence 16, Appl

ALIGNMENTS

RESULT 1
US-08-630-915A-197
; Sequence 197, Application US/08630915A
; Patent No. 6309820
; GENERAL INFORMATION:
; APPLICANT: SPARKS, Andrew B.
; APPLICANT: HOFFMAN, No. 6309820h
; APPLICANT: KAY, Brian K.
; APPLICANT: FOWLKES, Dana M.
; APPLICANT: MCCONNELL, Stephen J.
; TITLE OF INVENTION: POLYPEPTIDES HAVING A FUNCTIONAL
; TITLE OF INVENTION: DOMAIN OF INTEREST AND METHODS OF IDENTIFYING AND
; TITLE OF INVENTION: USING SAME
; NUMBER OF SEQUENCES: 227
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/630,915A
; FILING DATE: 03-APR-1996
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Miarock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 1101-174
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 197:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 971 bases
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA
; US-08-630-915A-197

Query Match 4.4%; Score 50; DB 3; Length 971;
Best Local Similarity 52.3%; Pred. No. 0.058;
Matches 134; Conservative 0; Mismatches 120; Indels 2; Gaps 1;

Db 2340 CTTCTCCGCGGCGCCTCCGCTCCGCGGACGCGCCCGCAGCTCCGCGCCCAACCAACAGTC 2281
Qy 709 CCAGTTCTGTGCTTCCGCGCCTACGAGACAGCCGCGCAGATGAGCTGTCCGTGCCCCG 768
Db 2280 CTTGCGCGGCGCTTACATAAACCGCGCA-CGCGGGGTCCGAGACCGCGCCCGCCGA 2222
Qy 769 GGGGCGCGGTGCGGTGTGGAACGTGAGACCGCGGTGTGCTATGCAAGGTACGG 828
Db 2221 GCGCGCGCGCGCGCGCGCTTACCTCAGCGCAGCGCGCGCGCTGTGCTCTCCG 2162
Qy 829 CGACCGCGGGGCTACTCTCCCGGTGTCTGCGCGCCGGAAGGCTGGCGCTCTCCT 888
Db 2161 CAGCGCGCGCGCTTCCGCGCAGTCTGCGCGCGCGCGCTTGTCTCTGTGGTT 2102
Qy 889 GAGCGGAGCGGGTTCCGTGGAGGAGACGACCGCGGGGTGAGGCGCGG--GCTTCCCT 946
Db 2101 GTGCGAGCGCGCGGTGCGCTGCGCTGCGCGCGCGAGCCCGCGCGCAACT 2042
Qy 947 GAACCTTCCAGGCGACCGCGCTTCCCGCCACCG-TGCCACCGGACCTTTCGCGGGCG 1005
Db 2041 GCGCGCTTGGCGCGCGGTGCGCGCGCGCGCGCGCGCGCGCGCGCTCAGCGCT 1982
Qy 1006 CATCCAGCGCTGTGCACTGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCA 1065
Db 1981 CAGCGCGCGCGCGCGCGCTTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1922
Qy 1066 GCGCGCGCTCGAGGTGTGCTGACTTCTGTGCGCGAC 1102
Db 1921 ACGCCAGGGGCGGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1885

RESULT 8
US-09-949-016-12765/c
; Sequence 12765, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12765
; LENGTH: 156894
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..-(156894)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12765

Query Match 4.4%; Score 49.4; DB 4; Length 156894;
Best Local Similarity 46.5%; Pred. No. 0.22;
Matches 296; Conservative 0; Mismatches 336; Indels 5; Gaps 4;
Qy 470 CACAGTCTGGAGGCTCAGAGCCTGCGCTGCGCTTGTATCCAGGACACGCGG 529
Db 2520 CGCAGCTCTGTCGCGCGCACTTGCCTTCCCGCGCGCGCGCGCGCGCGCGCGCG 2461
Qy 530 GATAGCGCTTTTCAGGCGGAGGAGAGCGCTGAGC-TGCTCTGCGGACCCCTC 588
Db 2460 ACTCGGCGCCTTCGAGGAGGAGCGCGGCGCGCGCGCGCGCGCGCGCGCGCGCG 2401

Qy 589 AGCTGTGTGCTGTGAGAAAGAACCGCGACAGCGCTGTGTTCAGCGCCCTACCT 648
Db 2400 GCGCGGAGCAGACGCGCGCAACCGAGGCTGTGCGAGCGGTGCGCGGCACTCG 2341
Qy 649 GAGGAGGCGGCGCGCGGCAAGCGCGGAGGAGCGCGCTCCTAGGAGACAGCGTCC 708
Db 2340 CTTCTCCGCGGCGCGCTTCCGCTCCCGGAGCGCGCGCGAGCTCCGCGCCCAACCAAGTC 2281
Qy 709 CCAGTTCTGTGCTTCCCGCGCTTACGAGACGACCGCGCAGATGAGTGTCCGTGCCCCG 768
Db 2280 CTTGCGCGGCGCTTACATAAACCGCGCA-CGCGGGGTCCGAGACCGCGCCCGCA 2222
Qy 769 GGGGCGCGGTGCGGTGTGGAACGTGAGACCGCGGTGTGCTATGCAAGGTACGG 828
Db 2221 GCGCGCGCGCGCGCGCGCTTACCTCAGCGCAGCGCGCGCGCGCTGTGCTCTCCG 2162
Qy 829 CGACCGCGGGGCTACTCTCCCGGTGTCTGCGCGCGCGGAAGGCTGGCGCTCTCCT 888
Db 2161 CAGCGCGCGCGCTTCCGCGCAGTCTGCGCGCGCGCGCTTGTCTCTGTGGTT 2102
Qy 889 GAGCGGAGCGGGTTCCGTGGAGGAGACGACCGCGGGGTGAGGCGCGG--GCTTCCCT 946
Db 2101 GTGCGAGCGCGCGGTGCGCTGCGCTGCGCGCGCGAGCCCGCGCGCAACT 2042
Qy 947 GAACCTTCCAGGCGACCGCGCTTCCCGCCACCG-TGCCACCGGACCTTTCGCGGGCG 1005
Db 2041 GCGCGCTTGGCGCGCGGTGCGCGCGCGCGCGCGCGCGCGCGCGCTCAGCGCT 1982
Qy 1006 CATCCAGCGCTGTGCACTGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCA 1065
Db 1981 CAGCGCGCGCGCGCGCGCTTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1922
Qy 1066 GCGCGCGCTCGAGGTGTGCTGACTTCTGTGCGCGAC 1102
Db 1921 ACGCCAGGGGCGGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1885

RESULT 9
US-09-949-016-12766/c
; Sequence 12766, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; SOFTWARE: FastSeq for Windows Version 4.0
; NUMBER OF SEQ ID NOS: 207012
; SEQ ID NO 12766
; LENGTH: 156894
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..-(156894)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12766

Query Match 4.4%; Score 49.4; DB 4; Length 156894;
Best Local Similarity 46.5%; Pred. No. 0.22;
Matches 296; Conservative 0; Mismatches 336; Indels 5; Gaps 4;
Qy 470 CACAGTCTGGAGGCTCAGAGCCTGCGCTGCGCTTGTATCCAGGACACGCGG 529
Db 2520 CGCAGCTCTGTCGCGCGCACTTGCCTTCCCGCGCGCGCGCGCGCGCGCGCGCG 2461

QY 530 GATAGGCTTTTTCAGGCGCAGCCCGAGAGAGCTTGGACG-TGCTGCTGGCGACCCCTC 588
DB 2460 ACTCGGCGCCCTCGGAGAGGGACGCGGGGCGCCCGCGGCGAGCTTCTTCGCGGAGCCGA 2401
QY 589 AGGCTGGTGGTGGTGGAGAACGAAGACCGGCGAGACCGCTTGGTTTCCAGCGCCCTACTC 648
DB 2400 GCGCGGAGCAGACGCGGCGAACCGAGAGGTGGTGGGAGGCGGTGCCCGGCGACCTCTG 2341
QY 649 GGAGAGGCGGCGCGCGGCGCAAGGCGGAGAGGCGCGCTCCCTTAGGAGCAGCGGTCC 708
DB 2340 CTCTCTCGGCGGCGCCCTCCGCTCCCGGAGACGGCGCGAGCTCCGGCGCCCAACCAAGTC 2281
QY 709 CCAGTTCTGTCTTCCCGGCTTACAGAGCAGCGCGCAGATGAGCTGCTCGTGCCCGC 768
DB 2280 CTTGCGCGGCGCTTACATAAACGCGCGCA-CCGCGGGGTCCGAGACCGCGCGCCCGCGCA 2222
QY 769 GGGGCGCGCGCTGCGCTGTGGAAACGTACAGACCGCGCTGGTGGCTATGACGATACGG 828
DB 2221 GCCGCGGCGCGCGCGCGCCCTCACCTCAGCGCAGCGCGCGCGCTGGTCTCTCG 2162
QY 829 CGACCGGCGGCGCTACTCCCGCGGTGCTGCTGCGGCGGAGAGGCTGGGCGCTCTCT 888
DB 2161 CACGCGCGCGCGCGCTTCCGCGCGAGTCTCGCGCGCGCGCGCTCTGCGCTCGTGGT 2102
QY 889 GAGCGGAGCGGGTTCGCTGGAGAGACACCGCGGCGGTGAGCGCCGGG--GCTTCCCT 946
DB 2101 GTGCGAGCGCGCGCGCTGGCGCTGCGCGCGCGCGCGCGCGCGCTCAGCGCT 2042
QY 1066 CATCAGAGCGCTGCTGACCGTACACGAGGCGCTTGGAGCGGCGCGCGCGCGCGCA 1065
DB 1981 CAGCGCGCGCGCGCGCTCGCGCGCGCTCCCGCGCGCGCGCGCGCGCGCGCGCGCG 1922
QY 1066 GGGCGCGCGCTCGAGGCTGGTGGACTCTGTGCGCGAC 1102
DB 1921 ACGCGAGGGGCGGGGCGCGCGAGCGCGCGGAC 1885

RESULT 10

US-09-949-016-16957/c
; Sequence 16957, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16957
; LENGTH: 156895
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(156895)
; OTHER INFORMATION: n = A,T,C or G

Query Match 4.4%; Score 49.4; DB 4; Length 156895;
Best Local Similarity 46.5%; Pred. No. 0.22;

Matches 296; Conservative 0; Mismatches 336; Indels 5; Gaps 4;

QY 470 CACAGTCTGGAGGCTCAGAGCCTGGCTGCTGCTGAGCCCTTCTGTACCCAGGACACGCGG 529
DB 2520 CGCAGCCTCGTCCCGGCGCACCTTCCCTCCCGGCGGCGAGCGCGGCGCATCCCCACGCGG 2461
QY 530 GATAGGCTTTTTCAGGCGCAGCCCGAGAGAGCTTGGACG-TGCTGCTGGCGACCCCTC 588
DB 2460 ACTCGGCGCCCTCGGAGGGACCGGGGCGCGCGGCGAGCTTCTTCGCGGAGCCGA 2401
QY 589 AGGCTGGTGGTGGTGGAGAACGAAGACCGGCGAGACCGCTTGGTTTCCAGCGCCCTACTC 648
DB 2400 GCGCGGAGCAGACGCGGCGAACCGAGAGGTGGTGGGAGGCGGTGCCCGGCGACCTCTG 2341
QY 649 GGAGAGGCGGCGCGGCGCGCAAGGCGGAGAGGCGCGCTCCCTTAGGAGCAGCGGTCC 708
DB 2340 CTTTCTCGCGGCGCCCTCCGCTCCCGGAGCGCGCGCTCGCGGCGCCCAACCAAGTC 2281
QY 709 CCAGTTCTGTCTTCCCGCGCTTACAGAGCAGCGCGCAGATGAGCTGCTCGTGCCCGC 768
DB 2280 CTTGCGCGGCGCTTACATAAACGCGCGCA-CCGCGGGGTCCGAGAGCGCGCGCGCGCA 2222
QY 769 GGGGCGCGCGCTGCGCGGTGGTGGAAACGTACAGACCGCGCGCTGGTGGCTATGACGATACGG 828
DB 2221 GCCGCGGCGCGCGCGCGCCCTCACCTCAGCGCAGCGCGCGCGCTGGTCTCTCG 2162
QY 829 CGACCGGCGGCGCTACTCCCGCGGTGCTGCTGCGGCGGAGAGGCTGGGCGCTCTCT 888
DB 2161 CACGCGCGCGCGCGCTTCCGCGCGAGTCTCGCGCGCGCGCGCTCTGCGCTCGTGGT 2102
QY 889 GAGCGGAGCGGGTTCGCTGGAGAGACACCGCGGCGGTGAGCGCCGGG--GCTTCCCT 946
DB 2101 GTGCGAGCGCGCGCGCTGGCGCTGCGCGCGCGCGCGCGCGCTCAGCGCT 2042
QY 947 GAACCTCTCCAGGCGACCGCGCTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1005
DB 2041 GCGCGCTTGGCGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 1982
QY 1006 CATCAGAGCGCTGCTGACCGTACACGAGGCGCTTGGAGCGGCGCGCGCGCGCGCGCA 1065
DB 1981 CAGCGCGCGCGCGCGCTCGCGCGCGCTCCCGCGCGCGCGCGCGCGCGCGCGCGCG 1922
QY 1066 GGGCGCGCGCTCGAGGCTGGTGGACTCTGTGCGCGAC 1102
DB 1921 ACGCGAGGGGCGGGGCGCGCGAGCGCGCGGAC 1885

RESULT 11

US-09-949-016-16958/c
; Sequence 16958, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16958
; LENGTH: 156895
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(156895)

; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-16958

Query Match 4.4%; Score 49.4; DB 4; Length 156895;
Best Local Similarity 46.5%; Pred. No. 0.22; Indels 5; Gaps 4;
Matches 296; Conservative 0; Mismatches 336; Indels 5; Gaps 4;
QY 470 CACAGTCTGGAGGCTCAGAGCCTGCGCTGCTGAGAGCCTTCTGTATCCAGGACACGCGG 529
Db 2520 CGCAGCCTCTCCCGGCGACCTTCCCTCCCGGCGAGCGCGGCATCCACGCGG 2461
QY 530 GATAGGCTTTTTCAGGCGGAGGCGGAGAGCCTTGACG-TGCTGCTGGGACACCCCTC 588
Db 2460 ACTCGGCGCTTCCGGAGGGAGCGCGGCGCGCTTCTCCCGGAGCGGA 2401
QY 589 AGCTGTGTGCTGTGGAGAACGAGACCGGAGAGCGCTGTGTTTCCAGCGCCCTACCT 648
Db 2400 GCCTCGGAGAGAGCGGCGGACCCAGAGGGTGGTGGAGGCGGTGCGCGGACCCCTCG 2341
QY 649 GGAGGAGCGGCGCGCGGCGGAGGAGGCGGCGGAGGAGGCGGCTTCTAGGAGAGCGGTC 708
Db 2340 CTTCTCGGCGGCGCTTACATAAACCGCGCA-CCGCGGGGTCCGAGCGCGCGCGCGCA 2281
QY 709 CGAGTCTGTGCTTCCCGCGCTTACAGAGAGCGGCGGAGATGAGCTGTCTGCTGCGCG 768
Db 2280 CTTCTCGGCGGCGCTTACATAAACCGCGCA-CCGCGGGGTCCGAGCGCGCGCGCGCA 2222
QY 769 GGGGCGCGGCTGCGGTGTGGAAACGTCAGACCGCGGCTGTGGCTATGCAGGTACGG 828
Db 2221 GCCGCGGCGCGCGCGCGCTTACCTAGCGGAGCGCGCGCGCGCTGTCTCTCG 2162
QY 829 CGACCGGCGGCGCTACTCTCCCGGCTGTGCTGCGCGCGGAGGAGGCGCTCTCTCT 888
Db 2161 CACGCGCGCGCGCTTCCGCGCGAGTCTCGCGCGCGGCGCGCTGTGCTGTGGTT 2102
QY 889 GAGCGGAGCGGGTTCCGTGGAGAGACGACCCGCGGCGGTGAGGCGCGG--GCTTCCCT 946
Db 2101 GTGCGAGCGGCGGCGGTGCGCTGCGCGCGCGCGCGCGGAGCGCGCGCGCAACT 2042
QY 947 GAACCTCTCCAGGCGGCGCGCTTCCCGCGCGGAGGAGGAGGAGGAGGAGGAGGAGG 1005
Db 2041 GCGCGCTTGGCGCGGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1982
QY 1006 CATCCAGAGCGCTGTGACCGCTCACAGCGAGGCGCTTGGAGCGCGCGCGCGCGCA 1065
Db 1981 CAGCGCGCGCGCGCGCGCTTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1922
QY 1066 GGGCGCGCTCGAGGTTGCTGAGTCTGTGCGGAC 1102
Db 1921 ACGCCGAGGGCGGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1885

RESULT 12

US-09-949-016-16959/c
; Sequence 16959, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT FILING DATE: 2000-04-14
; PRIOR FILING DATE: 2000-04-14
; PRIOR FILING DATE: 2000-10-20
; PRIOR FILING DATE: 2000-10-20
; PRIOR FILING DATE: 2000-10-03
; PRIOR FILING DATE: 2000-10-03
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 16959
; LENGTH: 156895

; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(156895)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-16959

Query Match 4.4%; Score 49.4; DB 4; Length 156895;
Best Local Similarity 46.5%; Pred. No. 0.22; Indels 5; Gaps 4;
Matches 296; Conservative 0; Mismatches 336; Indels 5; Gaps 4;
QY 470 CACAGTCTGGAGGCTCAGAGCCTGCGCTGCTGAGAGCCTTCTGTATCCAGGACACGCGG 529
Db 2520 CGCAGCCTCTCCCGGCGACCTTCCCTCCCGGCGAGCGCGGCATCCACGCGG 2461
QY 530 GATAGGCTTTTTCAGGCGGAGGCGGAGAGCCTTGACG-TGCTGCTGGGACACCCCTC 588
Db 2460 ACTCGGCGCTTCCGGAGGGAGCGCGGCGCGCTTCTCCCGGAGCGGA 2401
QY 589 AGCTGTGTGCTGTGGAGAACGAGACCGGAGAGCGCTGTGTTTCCAGCGCCCTACCT 648
Db 2400 GCCTCGGAGAGAGCGGCGGACCCAGAGGGTGGTGGAGGCGGTGCGCGGACCCCTCG 2341
QY 649 GGAGGAGCGGCGCGCGGCGGAGGAGGCGGCGGAGGAGGCGGCTTCTAGGAGAGCGGTC 708
Db 2340 CTTCTCGGCGGCGCTTACATAAACCGCGCA-CCGCGGGGTCCGAGCGCGCGCGCGCA 2281
QY 709 CGAGTCTGTGCTTCCCGCGCTTACAGAGAGCGGCGGAGATGAGCTGTCTGCTGCGCG 768
Db 2280 CTTCTCGGCGGCGCTTACATAAACCGCGCA-CCGCGGGGTCCGAGCGCGCGCGCGCA 2222
QY 769 GGGGCGCGGCTGCGGTGTGGAAACGTCAGACCGCGGCTGTGGCTATGCAGGTACGG 828
Db 2221 GCCGCGGCGCGCGCGCGCTTACCTAGCGGAGCGCGCGCGCGCTGTCTCTCG 2162
QY 829 CGACCGGCGGCGCTACTCTCCCGGCTGTGCTGCGCGCGGAGGAGGAGGAGGAGGAGG 888
Db 2161 CACGCGCGCGCGCTTCCGCGCGAGTCTCGCGCGCGGCGCGCTGTGCTGTGGTT 2102
QY 889 GAGCGGAGCGGGTTCCGTGGAGAGACGACCCGCGGCGGTGAGGCGCGG--GCTTCCCT 946
Db 2101 GTGCGAGCGGCGGCGGTGCGCTGCGCGCGCGCGCGCGGAGCGCGCGCGCAACT 2042
QY 947 GAACCTCTCCAGGCGGCGCGCTTCCCGCGCGGAGGAGGAGGAGGAGGAGGAGGAGG 1005
Db 2041 GCGCGCTTGGCGCGGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1982
QY 1006 CATCCAGAGCGCTGTGACCGCTCACAGCGAGGCGCTTGGAGCGCGCGCGCGCGCA 1065
Db 1981 CAGCGCGCGCGCGCGCGCTTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1922
QY 1066 GGGCGCGCTCGAGGTTGCTGAGTCTGTGCGGAC 1102
Db 1921 ACGCCGAGGGCGGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1885

RESULT 13

US-08-770-379-16
; Sequence 16, Application US/08770379
; Patent No. 5849564
; GENERAL INFORMATION:
; APPLICANT: Chang, Yuan
; APPLICANT: Bohenzky, Roy A.
; APPLICANT: Russo, James J.
; APPLICANT: Edelman, Isidore S.
; APPLICANT: Moore, Patrick S.
; TITLE OF INVENTION: POLYPEPTIDES FROM KAPOSI'S SARCOMA-ASSOCIATED
; TITLE OF INVENTION: HERPESVIRUS, DNA ENCODING SAME AND USES THEREOF
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas


```
QY 534 GGCCTTTTACGGCGCAGGCCCCAGAGAGCCTTGACCTGTGCTGTCGCGCACCCCTCAGGCT 593
D 213 CGCCAGCGCGCGCAGCGCCCCCGGGGCGCCGAGCCCCGAGCGGGGTACGGGGCT 272
QY 594 GGTGGCTTGTGGAGACGAGACCGGACAGACCCCTGCTTTCAGCGCCCTACTCTGAGG 653
D 273 AGGCCACGAATAATTTTTTTTCGGGCGGCCCCCGAACCTCTCTCGGCCCCCGGTCCC 332
QY 654 AGCGGCCCCCGGC-CAAGGCGCGGAGGAGGCGCCGTCCTTAGGGAGCAGCGCTCCCCAG 712
D 333 CGCGGCGCGCGCGCCCCCGGGGGGTAAACAGAGGGGGGGGGATGCGGCGCGGC 392
QY 713 TTCTGTGCTTCCGCGCTACAGAGACGCGCGCAGATGAGTGTCCGTGCGCGGG 772
D 393 GCGCGCGCGCGCGCGCTTGTCTTTCGTTTTCGCGGCGCCCCCGGGCGCGAG 452
QY 773 GCGCGGTGCGCGTGTGGAAAGTCAGACCGCGGCTGTGGCTATGACAGGTACGCGGAC 832
D 453 CGCGCGCGCGCGCGCGCCCCCTCCCGCGGGGGCTCGCGGGGGGGCCCCCTGTCCC 512
QY 833 CGCGCGCGGCTACTCCCGCGGTGCTGCTGCGGCGCGAGGGGCTGGCGCTCTCTGAGC 892
D 513 CGCGCGGCGCGGACCCCCCGCGCGCGCGCGCGCGCGATCCCGCGGGCGCGCCCC 572
QY 893 GGGACCGGGTTCCGTGGAGAGACGACCCGCGGGGTAGGCGCGGGCTTCCCTGAACCC 952
D 573 TGCGGGGACGCGCGCGGCTGCGCGGCTCCCGCGCGATGCGGCGCGCGCGCG 632
QY 953 TCCGAGCCACCGCCCCCTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1012
D 633 TCAGGCGCGCGCGCGCGCGCGCTGCTCCCGCGCGCGCGCGCGCGCGCGCGAG 692
QY 1013 AG 1014
D 693 CG 694
```

RESULT 15
US-09-230-371A-16
; Sequence 16, Application US/09230371A
; Patent No. 6348586
; GENERAL INFORMATION:
; APPLICANT: Chang, Yuan
; APPLICANT: Bohenzky, Roy A
; APPLICANT: Russo, James J
; APPLICANT: Edelman, Isidore S
; APPLICANT: Moore, Patrick S
; TITLE OF INVENTION: UNIQUE ASSOCIATED KAPOSI'S SARCOMA VIRUS SEQUENCES AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: 45185-G-PCT-US
; CURRENT APPLICATION NUMBER: US/09/230,371A
; CURRENT FILING DATE: 1999-11-17
; PRIOR APPLICATION NUMBER: PCT/US97/13346
; PRIOR FILING DATE: 1997-07-22
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 16
; LENGTH: 801
; TYPE: DNA
; ORGANISM: Kaposi's sarcoma-associated herpesvirus
US-09-230-371A-16

```
Query Match 4.3%; Score 49.2; DB 3; Length 801;
Best Local Similarity 44.1%; Pred.No.0.084;
Matches 292; Conservative 0; Mismatch 368; Indels 2; Gaps 2;

QY 354 GCTCTTCGACCGCAACCCCTGGACCTGGAGCCCGCGCTGCGCACCCCGGCGAGCGGGTGA 413
D 34 GCGCGCGCCCTCCCGCTCCCGCGCTGCGCTCCCGCGCTGCGCGCGCGCGCGCGCGCGGA 93
QY 414 TCGTCCCAACCCAGAGAGAGCCTCTTTTCGCGCTGCGGGCGCGCTCTCCATCACA 473
```

```
Db 94 GCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 153
QY 474 GTCTGAGGCTCAGAGCCTGCGCTGCGCTGCGAGCCCTTCTGTATCCACAGGACACCGCGGATA 533
D 154 GGAGGGGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 212
QY 534 GGCCTTTTTCAGGCGCAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 593
D 213 CGCCAGCGCGCGCGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 272
QY 594 GGTGGCTGTTGAGAACGAGACCGGACAGACCGCTGCTTTCAGCGCCCTACTCTGAGG 653
D 273 AGGCCACGAATAATTTTTTTTCGGGCGGCCCCCGAACCTCTCTCGGCCCCCGGTCCC 332
QY 654 AGCGGCCCCCGGC-CAAGGCGCGGAGGAGGCGCCGTCCTTAGGGAGCAGCGCTCCCCAG 712
D 333 CGCGGCGCGCGCGCCCCCGGGGGGTAAACAGAGGGGGGGGGATGCGGCGCGGC 392
QY 713 TTCTGTGCTTCCGCGCTACAGAGACGCGCGCAGATGAGTGTCCGTGCGCGGG 772
D 393 GCGCGCGCGCGCGCGCGCTTGTCTTTCGTTTTCGCGGCGCCCCCGGGCGCGAG 452
QY 773 GCGCGGTGCGCGTGTGGAAAGTCAGACCGCGGCTGTGGCTATGACAGGTACGCGGAC 832
D 453 CGCGCGCGCGCGCGCGCGCCCCCTCCCGCGGGGGCTCGCGGGGGGGCCCCCTGTCCC 512
QY 833 CGCGGCGGCGCTACTCCCGCGGTGCTGCTGCGGCGCGAGGGGCTGGCGCTCTCTGAGC 892
D 513 CGCGGCGGCGCGGACCCCCCGCGCGCGCGCGCGCGCGATCCCGCGGGCGCGCCCC 572
QY 893 GGGACCGGGTTCCGTGGAGAGACGACCCGCGGGGTAGGCGCGGGCTTCCCTGAACCC 952
D 573 TGCGGGGACGCGCGCGGCTGCGCGGCTCCCGCGCGATGCGGCGCGCGCGCGCGCG 632
QY 953 TCCGAGCCACCGCCCCCTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1012
D 633 TCAGGCGCGCGCGCGCGCGCGCTGCTCCCGCGCGCGCGCGCGCGCGCGCGAG 692
QY 1013 AG 1014
D 693 CG 694
```

Search completed: May 30, 2005, 23:21:52
Job time : 324 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 30, 2005, 15:46:02 ; Search time 973 Seconds
(without alignments)
6887.099 Million cell updates/sec

Title: US-10-621-113-3
Perfect score: 1132
Sequence: 1 agccatggcagggcccccgcagatcc 1132

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N Geneseq_16Dec04:.*
1: Geneseqn1980s:.*
2: Geneseqn1990s:.*
3: Geneseqn2000s:.*
4: Geneseqn2001s:.*
5: Geneseqn2001bs:.*
6: Geneseqn2002as:.*
7: Geneseqn2002bs:.*
8: Geneseqn2003as:.*
9: Geneseqn2003bs:.*
10: Geneseqn2003cs:.*
11: Geneseqn2003ds:.*
12: Geneseqn2004as:.*
13: Geneseqn2004bs:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1132	100.0	1132	12	ADJ57070 Coding se
2	1116	98.6	1129	12	ADJ57068 Coding se
3	1107	97.8	1147	12	ADJ57072 Coding se
4	1101.8	97.3	1520	13	ADQ85428 Human tum
5	1091	96.4	1144	12	ADJ57074 Coding se
6	798	70.5	1911	5	AAS89078 DNA encod
7	798	70.5	2881	5	AAS89434 DNA encod
8	342.4	30.2	2922	11	ADM02780 Human cdn
9	77.2	6.8	114955	2	AAX53491 Human ade
10	71.4	6.3	114955	2	AAX53491 Human ade
11	56.8	5.0	2781	5	ABV29190 Human pro
12	56.8	5.0	2781	5	ABV29190 Human pro
13	55.6	4.9	1850	13	ABD33012 Human can
14	55.6	4.9	2196	13	ACN42908 Human dia
15	55.6	4.9	2215	13	ABD33013 Human can
16	55.6	4.9	2255	13	ABD33018 Human can
17	55.6	4.9	2422	13	ACN42907 Human dia
18	55.6	4.9	2445	13	ABD33016 Human can
19	55.6	4.9	2448	13	ACN42906 Human dia
20	55.6	4.9	2453	13	ACN42913 Human dia

21	55.6	4.9	2470	13	ACN42912	Human dia
22	55.6	4.9	2579	13	ACN42911	Human dia
23	55.6	4.9	2580	9	AAL62035	Human cel
24	55.6	4.9	2622	13	ABD33019	Human can
25	55.6	4.9	2847	13	ACN42905	Human dia
26	55.6	4.9	2909	13	ABD33017	Human can
27	55.6	4.9	3004	13	ACN42910	Human dia
28	55.6	4.9	3121	13	ACN42909	Human dia
29	54.8	4.8	4667	12	ADQ22939	Human sof
30	54.8	4.8	84428	12	ADM45913	Streptomy
31	54	4.8	1574	3	AAZ48807	Rice inob
32	53.2	4.7	4545	10	ADE52577	Human SQV
33	53.2	4.7	4545	13	ABD33445	Human can
34	53.2	4.7	95914	13	ABD33444	Human can
35	52.8	4.7	2529	12	ADH18908	Human cel
36	52.8	4.7	2763	12	ADH18909	Human cel
37	52.8	4.7	110000	12	ADQ97138_3	Continuation (4 of
38	51.8	4.6	2156	5	AAS83572	DNA encod
39	51.8	4.6	2156	6	AAD24870	Human fib
40	51.8	4.6	2156	12	ADN03778	Antipsori
41	51.8	4.6	2200	2	AAQ11010	Fibulin C
42	51.8	4.6	2349	6	AAD24868	Human fib
43	51.8	4.6	2355	2	AAQ11008	Fibulin A
44	51.8	4.6	2359	5	AAS83573	DNA encod
45	51.8	4.6	2359	6	AAD24871	Human fib

ALIGNMENTS

RESULT 1
ADJ57070
ID ADJ57070 standard; cdna; 1132 BP.
XX
AC ADJ57070;
XX
DT 06-MAY-2004 (first entry)
XX
DE Coding sequence for p41Nox, a regulatory protein for Nox enzyme.
XX
KW Nox regulatory protein; p41Nox; NADH oxidase; cytosolic; chromosome 16;
XX human; gene; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 5..1120
FT /*tag= a
FT /product= "Human p41Nox"
XX
XX WO2004007689-A2.
XX
XX 22-JAN-2004.
XX
XX 16-JUL-2003; 2003WO-US022246.
XX
XX 16-JUL-2002; 2002US-0396170P.
PR 23-AUG-2002; 2002US-0405647P.
XX
XX (UYEM-) UNIV EMORY.
XX
XX Lambeth JD, Cheng G;
DR WPI; 2004-122927/12.
DR P-PSDB; ADJ57071.
XX
XX New regulatory protein for hydrogenated nicotinamide adenine dinucleotide
PT phosphate oxidase, useful for preparing a medicament for affecting
PT superoxide formation or regulating cellular proliferation in an animal or
PT a human.
XX
PS Claim 8; SEQ ID NO 3; 70pp; English.
XX

CC The present sequence is that of cDNA encoding a regulatory protein.
CC denoted p41Nox, for Nox (NADH oxidase) enzymes. The protein sequence of
CC p41Nox was used as query in database screenings, and a genomic sequence
CC was identified from human chromosome 16. PCR primers based on this
CC sequence were used in RT-PCR experiments to look for expression in a
CC series of normal and cancer cell lines. A 350 bp PCR product was detected
CC in RNA from human foetal kidney, liver, lung and T84 colon cancer cells.
CC RNA from T84 cells and human foetal liver was then used to amplify the
CC full-length p41Nox cDNA by PCR. 13 Unique clones were obtained.
CC Sequencing revealed 4 p41Nox splice variants, of which nucleotide and
CC protein sequences are provided ADJ57068-ADJ57075. p41Nox proteins
CC function in the regulation of cell growth and are therefore implicated in
CC diseases involving abnormal cell growth, such as cancer. They may also
CC function in innate immune mechanisms of epithelial tissue or other
CC barrier cells, and hence may be involved in diseases of diminished
CC ability to fight infections or inflammatory conditions. The p41Nox
CC regulatory proteins and nucleic acids are useful for preparing a
CC medicament for affecting superoxide formation or regulating cellular
CC proliferation in an animal or a human. The regulatory proteins are also
CC useful in drug development, e.g. screening for drugs that regulate the
CC biological activity of the proteins, or in assays that relate to
CC assessment of abnormal growth or cellular proliferation including cancer.
XX
SQ Sequence 1132 BP; 173 A; 399 C; 391 G; 169 T; 0 U; 0 Other;

Query Match 100.0%; Score 1132; DB 12; Length 1132;
Best Local Similarity 100.0%; Pred. No. 2.5e-213;
Matches 1132; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGCCATGGCAGGCCCCCGATACCCAGTTTTCAGTCAAGGGGAGCCCTGGTGAGATCAA 60
DB 1 AGCCATGGCAGGCCCCCGATACCCAGTTTTCAGTCAAGGGGAGCCCTGGTGAGATCAA 60
QY 61 GAGGCTCCAAACGTTTGGCTTCTCTGTGGCTGGTTCAGACGGCAGCGACACTTCGTGCG 120
DB 61 GAGGCTCCAAACGTTTGGCTTCTCTGTGGCTGGTTCAGACGGCAGCGACACTTCGTGCG 120
QY 121 CAGGAGTTGGGACGAATTCAGGCAGCTCAAGAGACCTCAAGAGACCTTCCCGGTGA 180
DB 121 CAGGAGTTGGGACGAATTCAGGCAGCTCAAGAGACCTCAAGAGACCTTCCCGGTGA 180
QY 181 GCGGGCTGTCTGGGAGACTGACCGGTTCTCCAAAGCTTCTCGATGACCACTGTT 240
DB 181 GCGGGCTGTCTGGGAGACTGACCGGTTCTCCAAAGCTTCTCGATGACCACTGTT 240
QY 241 GGGACGCTGGGGCGCAAGAGCCGGCTGGCGGCTGCGAGCTGTGTTGAAACCTATTTC 300
DB 241 GGGACGCTGGGGCGCAAGAGCCGGCTGGCGGCTGCGAGCTGTGTTGAAACCTATTTC 300
QY 301 TCGGAGGCTGTGGGAGCTGACAGCGGCTGGCAGCGGCGGCTGAGATCACTGGCTGTT 360
DB 301 TCGGAGGCTGTGGGAGCTGACAGCGGCTGGCAGCGGCGGCTGAGATCACTGGCTGTT 360
QY 361 CGCAGCGCAACCCCTGGACCTGGAGCCCGCTGCGCTGCACCCGGCAGCCGGTGATCTGCC 420
DB 361 CGCAGCGCAACCCCTGGACCTGGAGCCCGCTGCGCTGCACCCGGCAGCCGGTGATCTGCC 420
QY 421 CACCCACAGAGAGAGAGCTCTTTTCTGCGCTGCGGGCCGCTCTCCATCCACAGTCTGA 480
DB 421 CACCCACAGAGAGAGAGCTCTTTTCTGCGCTGCGGGCCGCTCTCCATCCACAGTCTGA 480
QY 481 GGCTCAGAGCCCTGGCTGCTGAGCGCTTCTGTATCCAGGACACGCGGGATAGGCTTTT 540
DB 481 GGCTCAGAGCCCTGGCTGCTGAGCGCTTCTGTATCCAGGACACGCGGGATAGGCTTTT 540
QY 541 TCAGGCGCAGGCCCGCAGGAGAGCTGGAGCTGTGCTGCGGGCACCCCTCAGGCTGGTGCT 600
DB 541 TCAGGCGCAGGCCCGCAGGAGAGCTGGAGCTGTGCTGCGGGCACCCCTCAGGCTGGTGCT 600
QY 601 GGTGGAGAACGAAGACCGGAGACCGCTGTGTTTCCAGCGCCCTTACCTGGAGGAGCGCGC 660
DB 601 GGTGGAGAACGAAGACCGGAGACCGCTGTGTTTCCAGCGCCCTTACCTGGAGGAGCGCGC 660

QY 661 CCCGGGCCAAGCGCGGAGGAGCGCCGCTCCTAGGAGAGCAGCGTCCCGCTTCTGTGC 720
DB 661 CCCGGGCCAAGCGCGGAGGAGCGCCGCTCCTAGGAGAGCAGCGTCCCGCTTCTGTGC 720
QY 721 TTCCCGGCTTACGAGAGACAGCCCGCAGATGAGCTGTCCGCGCGGGGCGCGCT 780
DB 721 TTCCCGGCTTACGAGAGACAGCCCGCAGATGAGCTGTCCGCGCGGGGCGCGCT 780
QY 781 GCGGTGTGTAAGAGCTCAGACCCGGGCTGGTGGCTATGACAGTACCGGAGCCGGCGG 840
DB 781 GCGGTGTGTAAGAGCTCAGACCCGGGCTGGTGGCTATGACAGTACCGGAGCCGGCGG 840
QY 841 CCTACTCCCGCGTGTCTGCTGCGCGGAGGGCTGGCGCTCTCTCTGAGCGGACGG 900
DB 841 CCTACTCCCGCGTGTCTGCTGCGCGGAGGGCTGGCGCTCTCTCTGAGCGGACGG 900
QY 901 GTTCCGTGGAGAGACGACCCCGCGGGGTGAGCCCGGGCTTCCCTGAACCTCCAGGC 960
DB 901 GTTCCGTGGAGAGACGACCCCGCGGGGTGAGCCCGGGCTTCCCTGAACCTCCAGGC 960
QY 961 CACGCGCCCTCCCGCAGCGCTGCGCACCCGACCTTCGCGGGGCGGCATCAGAGCGCTG 1020
DB 961 CACGCGCCCTCCCGCAGCGCTGCGCACCCGACCTTCGCGGGGCGGCATCAGAGCGCTG 1020
QY 1021 CTGCACGCTCACGAGCGGCGCTGGAGCGGCGCCACGCGCGCAGGCGCCCTCCAGG 1080
DB 1021 CTGCACGCTCACGAGCGGCGCTGGAGCGGCGCCACGCGCGCAGGCGCCCTCCAGG 1080
QY 1081 GTGCGTGGAGCTTCTGTGCGCACCCCGCAGCGAGCAGTGAGCGGAGGATCC 1132
DB 1081 GTGCGTGGAGCTTCTGTGCGCACCCCGCAGCGAGCAGTGAGCGGAGGATCC 1132
RESULT 2
ADJ57068
ID ADJ57068 standard; cDNA; 1129 BP.
XX
AC ADJ57068;
XX
DT 06-MAY-2004 (first entry)
XX
DE Coding sequence for p41Nox, a regulatory protein for Nox enzyme.
XX
KW Nox regulatory protein; p41Nox; NADH oxidase; cytostatic; chromosome 16;
KW human; gene; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
CDS 5..1117
FT /*tag= a
FT /product= "p41Nox"
XX
PN WO2004007689-A2.
XX
PD 22-JAN-2004.
XX
PP 16-JUL-2003; 2003MO-US022246.
XX
PR 16-JUL-2002; 2002US-0396170P.
PR 23-AUG-2002; 2002US-0405647P.
XX
PA (UYEM-) UNIV EMORY.
XX
PI Lambeth JD, Cheng G;
XX
DR WPI; 2004-122927/12.
P-PSDB; ADJ57069.
XX
XX
XX New regulatory protein for hydrogenated nicotinamide adenine dinucleotide
PT phosphate oxidase, useful for preparing a medicament for affecting
PT superoxide formation or regulating cellular proliferation in an animal or
PT a human.

XX	Claim 8; SEQ ID NO 1; 70pp; English.	
PS	The present sequence is that of cDNA encoding a regulatory protein, denoted p41Nox, for Nox (NADH oxidase) enzymes. The protein sequence of p47phox was used as query in database screenings, and a genomic sequence was identified from human chromosome 16. PCR primers based on this sequence were used in RT-PCR experiments to look for expression in a series of normal and cancer cell lines. A 350 bp PCR product was detected in RNA from human foetal kidney, liver, lung and T84 colon cancer cells.	
CC	RNA from T84 cells and human foetal liver was then used to amplify the full-length p41Nox cDNA by PCR. 13 Unique clones were obtained.	
CC	Sequencing revealed 4 p41Nox splice variants, of which nucleotide and protein sequences are provided ADJ57068-ADJ57075. p41Nox proteins function in the regulation of cell growth and are therefore implicated in diseases involving abnormal cell growth, such as cancer. They may also function in innate immune mechanisms of epithelial tissue or other barrier cells, and hence may be involved in diseases of diminished ability to fight infections or inflammatory conditions. The p41Nox regulatory proteins and nucleic acids are useful for preparing a medicament for affecting superoxide formation or regulating cellular proliferation in an animal or a human. The regulatory proteins are also useful in drug development, e.g. screening for drugs that regulate the biological activity of the proteins, or in assays that relate to assessment of abnormal growth or cellular proliferation including cancer.	
XX	Sequence 1129 BP; 171 A; 399 C; 390 G; 169 T; 0 U; 0 Other;	
SQ	Query Match 98.6%; Score 1116; DB 12; Length 1129; Best Local Similarity 99.7%; Pred. No. 3.5e-210; Matches 1129; Conservative 0; Mismatches 0; Indels 3; Gaps 1;	
QY	1 AGCCATGGCAGGCCCCCGATACCCAGTTTCAGTGAAGGGGAGCCCTGGTGAGATCAA 60	Db 598 GGTGGAGAACGAAGACCGGCAGACCGCCTGGTTTCCAGGCGCCCTACCTGGAGAGCGGC 657
Db	1 AGCCATGGCAGGCCCCCGATACCCAGTTTCAGTGAAGGGGAGCCCTGGTGAGATCAA 60	QY 661 CCCGGGCCAAGGCCGGGAGGAGCCGCTCCCTAGGAGCAGCGGTCCCCAGTTCTGTGC 720
QY	61 GAGGCTCCAAACGTTTGGCTTCTCTGTGGCTGGTCAAGCGGAGCACCTTCGTGGC 120	Db 658 CCCGGGCCAAGGCCGGGAGGAGCCGCTCCCTAGGAGCAGCGGTCCCCAGTTCTGTGC 717
Db	61 GAGGCTCCAAACGTTTGGCTTCTCTGTGGCTGGTCAAGCGGAGCACCTTCGTGGC 120	QY 721 TTCCCGCGCTACGAGAGCAGCGGCAGATGAGTGTCCGTGCCCGCGGGCGCGCT 780
QY	121 CAGGAGTTGGGAGGAATTCAGGAGCTCAAGAGACCTTCAGGAGACCTTCCCGTGA 180	Db 718 TTCCCGCGCTACGAGAGCAGCGGCAGATGAGTGTCCGTGCCCGCGGGCGCGCT 777
Db	121 CAGGAGTTGGGAGGAATTCAGGAGCTCAAGAGACCTTCAGGAGACCTTCCCGTGA 177	QY 781 GCGCGTTTGGAAACGTCAAGACCGCGGCTGGTGGCTATGCAAGTACGGCGACCGGGCGG 840
QY	181 GCGGCGCTGCTGCGGAGATCTGACCGCGTTCCTCCAAAGCTTCTCGATGCAACCTGTT 240	Db 778 GCGCGTTTGGAAACGTCAAGACCGCGGCTGGTGGCTATGCAAGTACGGCGACCGGGCGG 837
Db	178 GCGGCGCTGCTGCGGAGATCTGACCGCGTTCCTCCAAAGCTTCTCGATGCAACCTGTT 237	QY 841 CTCTACTCCCGCGGTGCTCTGCGGCCGGAAGGGCTGGCGGTCTCTCTGAGCGGGACGGG 900
QY	241 GGGACCGCTGGGGCGCACGAGCGCGCGCTGGCGCGCTGAGCTGTTGGAACCTATTTC 300	Db 838 CTCTACTCCCGCGGTGCTCTGCGGCCGGAAGGGCTGGCGGTCTCTCTGAGCGGGACGGG 897
Db	238 GGGACCGCTGGGGCGCACGAGCGCGCGCTGGCGCGCTGAGCTGTTGGAACCTATTTC 297	QY 901 GTTCCGTGAGGAGAGACACCCCGCGGGTGAGGCCCGGGGCTTCCCTGAAACCTCCAGGC 960
QY	301 TCGGAGGCTGCTGGGACCTGCAGAGCGCTGGCAGCGGACCATCACTGGCTTCTT 360	Db 898 GTTCCGTGAGGAGAGACACCCCGCGGGTGAGGCCCGGGGCTTCCCTGAAACCTCCAGGC 957
Db	298 TCGGAGGCTGCTGGGACCTGCAGAGCGCTGGCAGCGGACCATCACTGGCTTCTT 357	QY 961 CACCGCCCTCCCGCCACCGTGCACCCGACCTTGGCCGGGGCCCATCCAGAGCGCGCTG 1020
QY	361 CGCACCGCAACCCCTGAGCTGAGCGCCGCGCTGCCACCCGCGAGCGCGGTGATCTCTGCC 420	Db 958 CACCGCCCTCCCGCCACCGTGCACCCGACCTTGGCCGGGGCCCATCCAGAGCGCGCTG 1017
Db	358 CGCACCGCAACCCCTGAGCTGAGCGCCGCGCTGCCACCCGCGAGCGCGGTGATCTCTGCC 417	QY 1021 CTGCACCGTCAACGAGCGGCTCTGAGCGGCGCCCGCCAGCGCGCGCGCTCCAGG 1080
QY	421 CACCCAGAGGAGACAGCCTTTCTTCTCGCGCTGCGGCGCGCTCTCCATCCACAGTCTGA 480	Db 1018 CTGCACCGTCAACGAGCGGCTCTGAGCGGCGCCCGCCAGCGCGCGCGCTCCAGG 1077
Db	418 CACCCAGAGGAGACAGCCTTTCTTCTCGCGCTGCGGCGCGCTCTCCATCCACAGTCTGA 477	QY 1081 GTGCGTGGACTCTGTGCGCGCACCCACGACGAGGAGCAGTGAGCGCGAGGATCC 1132
QY	481 GGCTCAGAGCTGCGCTGCTGCGGCTTCTGTACCCAGGACACCGGGATAGGCTTT 540	Db 1078 GTGCGTGGACTCTGTGCGCGCACCCACGACGAGCAGTGAGCGCGAGGATCC 1129
Db	478 GGCTCAGAGCTGCGCTGCTGCGGCTTCTGTACCCAGGACACCGGGATAGGCTTT 537	RESULT 3
QY	541 TCAGGCGCAGGCGCCAGAGAGCTGACAGCTGTGCTGCGGCGCCCTCCAGGCTGGTGGCT 600	ADJ57072
Db	538 TCAGGCGCAGGCGCCAGAGAGCTGACAGCTGTGCTGCGGCGCCCTCCAGGCTGGTGGCT 597	ID ADJ57072 standard; cDNA; 1147 BP.
QY	601 GGTGGAGAACGAAGACCGGCAGACCGCCTGGTTTCCAGCGCCCTACCTGGAGGAGCGGC 660	XX ADJ57072;
		AC ADJ57072;
		XX 06-MAY-2004 (first entry)
		DT Coding sequence for p41Nox, a regulatory protein for Nox enzyme.
		XX Nox regulatory protein; p41Nox; NADH oxidase; cytosolic; chromosome 16;
		XX human; gene; ss.
		OS Homo sapiens.
		XX Key Location/Qualifiers
		FT CDS 5..1135
		FT /*tag= a
		FT /product= "Human p41Nox"
		XX WO2004007689-A2.
		PN 22-JAN-2004.
		XX 16-JUL-2003; 2003WO-US022246.
		XX 16-JUL-2002; 2002US-0396170P.
		PR 23-AUG-2002; 2002US-0405647P.
		XX (UYEM-) UNIV EMORY.
		PA Lambeth JD, Cheng G;
		XX WPI; 2004-122927/12.
		DR P-PSDB; ADJ57073.
		XX New regulatory protein for hydrogenated nicotinamide adenine dinucleotide
		PT

XX	CC	New nucleic acid molecule and encoded polypeptide, for diagnosing, preventing or treating cell proliferative disorders such as cancer.
PT	CC	Claim 1; SEQ ID NO 2242; 5504pp; English.
XX	CC	The present invention describes an isolated tumour-associated antigenic target (TAT) nucleic acid comprising: (a) any of 4622 nucleotide sequences (see SEQ ID NO:1 to 4622); (b) the full-length coding region of (a); (c) the complement of (a) or (b); (d) a sequence that has 80% sequence identity to (a)-(c); or (e) a sequence that hybridises to (a)-(c). Also described: (1) an expression vector comprising the above nucleic acid; (2) a host cell comprising the above expression vector; (3) a process for producing a polypeptide; (4) an isolated polypeptide comprising: (a) an amino acid sequence encoded by any of the above nucleotide sequences; (b) an amino acid sequence encoded by the full-length coding region of the above nucleotide sequences; or (c) a sequence having at least 80% identical to (a) or (b); (5) a chimeric polypeptide comprising the above polypeptide fused to a heterologous polypeptide; (6) an isolated antibody that binds to the above polypeptide; (7) a process for producing the antibody; (8) an isolated oligopeptide that binds to the above polypeptide; (9) a tumour-associated antigenic target (TAT) binding organic molecule that binds to the above polypeptide; (10) a composition of matter comprising the above (chimeric) polypeptide, antibody, oligopeptide or TAT binding organic molecule, in combination with a carrier; (11) an article of manufacture comprising a container and the composition of matter contained within the container; (12) methods of inhibiting the growth of a cell that expresses the above protein, where the growth of the cell is at least in part dependent upon a growth potentiating effect of the above protein; (13) a method of therapeutically treating a mammal having a cancerous tumour comprising cells that express the above protein; (14) a method of determining the presence of a protein in a sample suspected of containing the protein described above; (15) methods of diagnosing the presence of a tumour in a mammal; (16) a method for treating or preventing a cell proliferative disorder associated with increased expression or activity of the above protein; and (17) a method of binding an antibody, oligopeptide or organic molecule to a cell that expresses the protein described above. The TAT sequences have cytostatic activities, and can be used in gene therapy. The composition and methods are useful for diagnosing, preventing or treating cancer. The composition is also used for preparing a medicament for the therapeutic treatment or diagnostic detection of a cell proliferative disorder or cancer. The present sequence represents a human TAT cDNA sequence from the present invention.
XX	CC	Sequence 1520 BP; 243 A; 512 C; 520 G; 245 T; 0 U; 0 Other;
QQ	CC	Query Match 97.3%; Score 1101.8; DB 13; Length 1520;
DB	CC	Best Local Similarity 99.5%; Pred. No. 2.2e-207;
XX	CC	Matches 1127; Conservative 0; Mismatches 2; Indels 4; Gaps 2
QY	CC	1 AGCCATGGCAGGCCCCGATACCCAGTTTCAGTCAAGGGCAGCCCTTGTCAGATCAA 60
DB	CC	269 AGCCATGGCAGGCCCCGATACCCAGTTTCAGTCAAGGGCAGCCCTTGTCAGATCAA 328
QY	CC	61 GAGGCTCCAAACGTTTGCTTCTCTGTGCGCTGGTTCAGACGGCAGCGACACCTTCGTGCG 120
DB	CC	329 GAGGCTCCAAACGTTTGCTTCTCTGTGCGCTGGTTCAGACGGCAGCGACACCTTCGTGCG 388
QY	CC	121 CAGGAGTTGGAGCGAAATTCAGGCGAGCTCAAGAGACCTCTCAAGGAGACCTTCCCGGTGGA 180
DB	CC	389 CAGGAGTTGGAGCGAAATTCAGGCGAGCTC--NAGACCCTCAAGGAGACCTTCCCGGTGGA 445
QY	CC	181 GCGGGGCTGCTGCGGAGATCTGACCGGTTCTCCCAAGCTTCTCGATGACCACTGTT 240
DB	CC	446 GCGGGGCTGCTGCGGAGATCTGACCGGTTCTCCCAAGCTTCTCGATGACCACTGTT 505
QY	CC	241 GGGACGGTGGGGCGCAGACCCGCGCTGCGCGCTGCGAGCTGTGTGGAAACCTATTTC 300
DB	CC	506 GGGACGGTGGGGCGCAGACCCGCGCTGCGCGCTGCGAGCTGTGTGGAAACCTATTTC 565
QY	CC	301 TCGGAGGCTGTGTCGCGACTGCAGAGCGGTGGCAGCGAGCCGACGATCACTGGGCTTCTT 360

```

FT XX /product= "Human p41Nox"
PN XX
XX WO2004007689-A2.
PD XX
XX 22-JAN-2004.
XX
XX 16-JUL-2003; 2003WO-US022246.
XX
XX 16-JUL-2002; 2002US-0396170P.
PR XX
XX 23-AUG-2002; 2002US-0405647P.
XX
XX (UYEM-) UNIV EMORY.
XX
XX Lambeth JD, Cheng G;
PI XX
XX WPI; 2004-122927/12.
DR XX
XX P-PSDB; ADJ57075.
XX
XX New regulatory protein for hydrogenated nicotinamide adenine dinucleotide
PT phosphate oxidase, useful for preparing a medicament for affecting
PT superoxide formation or regulating cellular proliferation in an animal or
PT a human.
XX
XX Claim 8; SEQ ID NO 7; 70pp; English.
XX
XX The present sequence is that of cDNA encoding a regulatory protein,
CC denoted p41Nox, for Nox (NADH oxidase) enzymes. The protein sequence of
CC p47phox was used as query in database screenings, and a genomic sequence
CC was identified from human chromosome 16. PCR primers based on this
CC sequence were used in RT-PCR experiments to look for expression in a
CC series of normal and cancer cell lines. A 350 bp PCR product was detected
CC in RNA from human foetal kidney, liver, lung and T84 colon cancer cells.
CC RNA from T84 cells and human foetal liver was then used to amplify the
CC full-length p41Nox cDNA by PCR. 13 Unique clones were obtained.
CC Sequencing revealed 4 p41Nox splice variants, of which nucleotide and
CC protein sequences are provided ADJ57068-ADJ57075. p41Nox proteins
CC function in the regulation of cell growth and are therefore implicated in
CC diseases involving abnormal cell growth, such as cancer. They may also
CC function in innate immune mechanisms of epithelial tissue or other
CC barrier cells, and hence may be involved in diseases of diminished
CC ability to fight infections or inflammatory conditions. The p41Nox
CC regulatory proteins and nucleic acids are useful for preparing a
CC medicament for affecting superoxide formation or regulating cellular
CC proliferation in an animal or a human. The regulatory proteins are also
CC useful in drug development, e.g. screening for drugs that regulate the
CC biological activity of the proteins, or in assays that relate to
CC assessment of abnormal growth or cellular proliferation including cancer.
XX
XX Sequence 1144 BP; 173 A; 404 C; 396 G; 171 T; 0 U; 0 Other;
SQ

```

```

Query Match 96.4%; Score 1091; DB 12; Length 1144;
Best Local Similarity 98.4%; Pred. No. 2.9e-205;
Matches 1129; Conservative 0; Mismatches 0; Indels 18; Gaps 2;

QY 1 AGCCATGTCAGCGCCCGATACCCAGTTTCAGTGCAGAGGGCAGCCCTGGTGACAGATCAA 60
DB 1 AGCCATGTCAGCGCCCGCCGATACCCAGTTTCAGTGCAGAGGGCAGCCCTGGTGACAGATCAA 60
QY 61 GAGGCTCCAAACGTTTGCTTCTCTGTGCGTGTGTGTCAGAGCGGACACCTTCGTGCG 120
DB 61 GAGGCTCCAAACGTTTGCTTCTCTGTGCGTGTGTGTCAGAGCGGACACCTTCGTGCG 120
QY 121 CAGGAGTTGGACGATTCAGGACGCTCAAGAACCCCTCAAGAGACCTTCGCGTGGGA 180
DB 121 CAGGAGTTGGACGATTCAGGACGCTCAAGAACCCCTCAAGAGACCTTCGCGTGGGA 177
QY 181 GCGCGCTGCTGCGGAGATCTGACCGGTTCTCCAAAGCTTCTC----- 226
DB 178 GCGCGGCTGCTGCGGAGATCTGACCGGTTCTCCAAAGCTTCTCGGTACGGCAGCGCT 237
QY 227 -GATGCAACCATGTTGGAGACGCGTGGGCGCACAGCGCGCGCTGGCGCGCTGCACT 285
DB 238 GGATGCAACCATGTTGGAGACGCGTGGGCGCACAGCGCGCGCTGGCGCGCTGCACT 297

```

```

QY 286 GTTGGAAACCTATTCTCGAGGCTGCTGGCGACTGACAGAGCGCTGGCAACCGGCGCAC 345
DB 298 GTTGGAAACCTATTCTCGAGGCTGCTGGCGACTGACAGAGCGCTGGCAACCGGCGCAC 357
QY 346 GATCACTGGCTTCTTTCGCAACCGCCCTGGAGCTGAGAGCCCGCTGCCACCGCGGAG 405
DB 358 GATCACTGGCTTCTTTCGCAACCGCCCTGGAGCTGAGAGCCCGCTGCCACCGCGGAG 417
QY 406 CCGGCTGATCTGCGCCACCGCCAGAGGAGCAGCTTCTTCTCGCGCTGGCGGCGCGCTCTC 465
DB 418 CCGGCTGATCTGCGCCACCGCCAGAGGAGCAGCTTCTTCTCGCGCTGGCGGCGCGCTCTC 477
QY 466 CATCCACAGTCTGAGGCTCAGAGCTGCGCTGCGCTGAGCCCTTCTGTACCCAGGACAC 525
DB 478 CATCCACAGTCTGAGGCTCAGAGCTGCGCTGCGCTGAGCCCTTCTGTACCCAGGACAC 537
QY 526 GCGGGATAGGCTTTTTCAGGCGCAGAGCCCTGAGAGAGCTGCTGCTGCTGCGGACACC 585
DB 538 GCGGGATAGGCTTTTTCAGGCGCAGAGCCCTGAGAGAGCTGCTGCTGCTGCTGCGGACACC 597
QY 586 CTCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 645
DB 598 CTCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 657
QY 646 CCTGGAGAGCGCGCCCGCGGCCAAGCGCGGAGGAGGCGCGCTCCCTAGGAGGAGCGG 705
DB 658 CCTGGAGAGCGCGCCCGCGGCCAAGCGCGGAGGAGGCGCGCTCCCTAGGAGGAGCGG 717
QY 706 TCCCGAGTCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 765
DB 718 TCCCGAGTCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 777
QY 766 GCGGGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 825
DB 778 GCGGGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 837
QY 826 GCGGACCGCGCGCGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 885
DB 838 GCGGACCGCGCGCGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 897
QY 886 CCTGAGCGGACCGGGTTCGCTGAGAGAGACGACCCCGCGGCTGAGGCGCGGGGCTTCCC 945
DB 898 CCTGAGCGGACCGGGTTCGCTGAGAGAGACGACCCCGCGGCTGAGGCGCGGGGCTTCCC 957
QY 946 TGAACCTCTCCAGGACACCGCGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1005
DB 958 TGAACCTCTCCAGGACACCGCGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1017
QY 1006 CATCCAGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1065
DB 1018 CATCCAGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1077
QY 1066 GGGCGCGCTTCTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1125
DB 1078 GGGCGCGCTTCTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1137
QY 1126 AGGATCC 1132
DB 1138 AGGATCC 1144

```

```

RESULT 6
AAS89078
ID AAS89078 standard; cDNA; 1911 BP.
XX
AC AAS89078;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #24882.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;

```

food supplement; medical imaging; diagnostic; genetic disorder; ss.

Homo sapiens.

WO200175067-A2.

11-OCT-2001.

30-MAR-2001; 2001WO-US0008631.

31-MAR-2000; 2000US-00540217.

23-AUG-2000; 2000US-00649167.

(HYSE-) HYSEQ INC.

Drmanac RT, Liu C, Tang YT;

WPI; 2001-639362/73.

P-PSDB; ABG24891.

New isolated polynucleotide and encoded polypeptides, useful in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits and to assess biodiversity.

Claim 1; SEQ ID NO 24882; 103pp; English.

The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic coding sequences of the invention. Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

Sequence 1911 BP; 397 A; 545 C; 587 G; 382 T; 0 U; 0 Other;

Query Match 70.5%; Score 798; DB 5; Length 1911;

Best Local Similarity 98.2%; Pred. No. 1.2e-147;

Matches 823; Conservative 0; Mismatches 0; Indels 15; Gaps 1;

5 ATGGCAGGCCCCGATACCCAGTTTCAGTGCAGGGGCGAGCCCTGGTGCAGATCAAGAGG 64

1 ATGGCAGGCCCCGATACCCAGTTTCAGTGCAGGGGCGAGCCCTGGTGCAGATCAAGAGG 60

65 CTCGAAACGTTTGCCTTCTGTGCGCTGGTGCAGCGGAGGACACCTTCTGTGCGCAGG 124

61 CTCGAAACGTTTGCCTTCTGTGCGCTGGTGCAGCGGAGGACACCTTCTGTGCGCAGG 120

125 AGTTGGGAGCAATTCAGGAGGCTCAAGAGACCTCAAGGAGACCTTCCGGTGGAGGCG 184

121 AGTTGGGAGCAATTCAGGAGGCTCAAGAGACCTCAAGGAGACCTTCCGGTGGAGGCG 180

185 GGCCTGCTCGGAGATCTGACCGCGTTCTCCCAAGCTTCTC-----GAT 229

181 GGCCTGCTCGGAGATCTGACCGCGTTCTCCCAAGCTTCTCGGTACAGCCAGCTGGAT 240

230 GCACCACTTTGGAGCGCTGGGGCGCAAGCCGCGCTGGCGCGCTGAGCTGTGTTG 289

|||||

|||||

|||||

|||||

Db 241 GCACCACTTTGGAGCGCTGGGGCGCACGAGCCGGCGCTGCGCGCTGCGAGCTGTG 300

QY 290 GAAACCTATTTCGGAGGCTGTGGCGACTGACGAGCGGTGCGACGAGCCGAGATC 349

Db 301 GAAACCTATTTCGGAGGCTGTGGCGACTGACGAGCGGTGCGACGAGCCGAGATC 360

QY 350 ACTGGCTTTTTCGACCCGCAACCCCTGGACCTGGAGCCCGCGCTGCCACCCCGCAGCCGG 409

Db 361 ACTGGCTTTTTCGACCCGCAACCCCTGGACCTGGAGCCCGCGCTGCCACCCCGCAGCCGG 420

QY 410 GTGATCTTCCGACCCCGCAGAGGAGCCTCTTTCTGCGCGCTGCGGCGCGCTTCTCATC 469

Db 421 GTGATCTTCCGACCCCGCAGAGGAGCCTCTTTCTGCGCGCTGCGGCGCGCTTCTCATC 480

QY 470 CACAGTCTGGAGGCTCAGAGCCTGCGCTGCGAGCCCTTCTGTATCCCGAGACACGCGG 529

Db 481 CACAGTCTGGAGGCTCAGAGCCTGCGCTGCGAGCCCTTCTGTATCCCGAGACACGCGG 540

QY 530 GATAGGCTTTTTCAGGCGCAGGCCCGCAGGAGACCTGGAGCGTGTGCTGCGGCAACCCCTCA 589

Db 541 GATAGGCTTTTTCAGGCGCAGGCCCGCAGGAGACCTGGAGCGTGTGCTGCGGCAACCCCTCA 600

QY 590 GCGTGTGCTGCTGGAGAACGAGACCGGACAGCCGCTGCTTCCAGGCGCCCTACCTG 649

Db 601 GCGTGTGCTGCTGGAGAACGAGACCGGACAGCCGCTGCTTCCAGGCGCCCTACCTG 660

QY 650 GAGGAGCGCGCCCGGCGCAAGCGCGGAGGAGCGCCGCTCCCTAGGGAGCAGCGTCCC 709

Db 661 GAGGAGCGCGCCCGGCGCAAGCGCGGAGGAGCGCCGCTCCCTAGGGAGCAGCGTCCC 720

QY 710 CAGTTCTGTGCTTCCGCGCTTACGAGACGACCGCGCAGATGAGCTGTCCGTGCGCGG 769

Db 721 CAGTTCTGTGCTTCCGCGCTTACGAGACGACCGCGCAGATGAGCTGTCCGTGCGCGG 780

QY 770 GGGGCGCGCTGGCGGTGTGGAACCTCAGACCGCGCGCTGGCTATCAGGTACG 827

Db 781 GGGGCGCGCTGGCGGTGTGGAACCTCAGACCGCGCGCTGGCTATCAGGTACG 838

RESULT 7

AAS89434

ID AAS89434 standard; cDNA; 2881 BP.

AC AAS89434;

DT 13-FEB-2002 (first entry)

DE DNA encoding novel human diagnostic protein #25238.

Human; chromosome mapping; gene mapping; gene therapy; forensic;

food supplement; medical imaging; diagnostic; genetic disorder; ss.

OS Homo sapiens.

XX WO200175067-A2.

XX 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US0008631.

XX 31-MAR-2000; 2000US-00540217.

XX 23-AUG-2000; 2000US-00649167.

XX (HYSE-) HYSEQ INC.

XX Drmanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.

XX P-PSDB; ABG25247.

XX New isolated polynucleotide and encoded polypeptides, useful in

diagnostics, forensics, gene mapping, identification of mutations

responsible for genetic disorders or other traits and to assess

PT biodiversity.
PS Claim 1; SEQ ID NO 25238; 103pp; English.
XX The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, CC and in recombinant production of (II). The polynucleotides are also used CC in diagnostics as expressed sequence tags for identifying expressed CC genes. (I) is useful in gene therapy techniques to restore normal CC activity of (II) or to treat disease states involving (II). (II) is CC useful for generating antibodies against it, detecting or quantitating a CC polypeptide in tissue, as molecular weight markers and as a food CC supplement. (II) and its binding partners are useful in medical imaging CC of sites expressing (II). (I) and (II) are useful for treating disorders CC involving aberrant protein expression or biological activity. The CC polypeptide and polynucleotide sequences have applications in CC diagnostics, forensics, gene mapping, identification of mutations CC and responsible for genetic disorders or other traits to assess biodiversity CC and to produce other types of data and products dependent on DNA and CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic CC coding sequences of the invention. Note: The sequence data for this CC patent did not appear in the printed specification, but was obtained in CC electronic format directly from WIPO at CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 2881 BP; 643 A; 803 C; 810 G; 625 T; 0 U; 0 Other;
Query Match 70.5%; Score 798; DB 5; Length 2881;
Best Local Similarity 98.2%; Pred. No. 1.2e-147;
Matches 823; Conservative 0; Mismatches 0; Indels 15; Gaps 1;
QY 5 ATGGCAGGCCCCGATACCCAGTTTCAGTGCAGGGGAGCCCTGTCAGATCAGAGG 64
DB 1 ATGGCAGGCCCCGATACCCAGTTTCAGTGCAGGGGAGCCCTGTCAGATCAGAGG 60
QY 65 CTCGAAACGTTTGCTTCTCTGTGGCTGGTTCAGACGCGCAGCACCTTCGTGGCAGG 124
DB 61 CTCGAAACGTTTGCTTCTCTGTGGCTGGTTCAGACGCGCAGCACCTTCGTGGCAGG 120
QY 125 AGTTGGGACGAATTTCAGGACGCTCAAGAGACCCCTCAAGGAGACCTTCCGGTGGAGCG 184
DB 121 AGTTGGGACGAATTTCAGGACGCTCAAGAGACCCCTCAAGGAGACCTTCCGGTGGAGCG 180
QY 185 GGCCTGCTCGGAGATCTGACGGCGTTCTCCAAAGCTTCTC-----GAT 229
DB 181 GGCCTGCTCGGAGATCTGACGGCGTTCTCCAAAGCTTCTCGGTTCAGGCCAGCGTGGAT 240
QY 230 GCACCACTGTTGGGACGCTGGGGGCGACGAGCCGCGCTGGCGCGCTGCAGCTGTTG 289
DB 241 GCACCACTGTTGGGACGCTGGGGGCGACGAGCCGCGCTGGCGCGCTGCAGCTGTTG 300
QY 290 GAAACCTATTCTCGGAGGCTGCTGGCGACTGCGAGCGCGTGGCAGGAGCCGACGATC 349
DB 301 GAAACCTATTCTCGGAGGCTGCTGGCGACTGCGAGCGCGTGGCAGGAGCCGACGATC 360
QY 350 ACTGGCTTCTTCGACCCGACCCCTGACCTGAGGCGCGCGCTGCGACCCGCGAGCGCG 409
DB 361 ACTGGCTTCTTCGACCCGACCCCTGACCTGAGGCGCGCGCTGCCACCCGCGAGCGCG 420
QY 410 GTGATCTCTCCACCCAGAGGAGCAGCTCTTTCTCGCGCTGGGGCGCGCTCTCCATC 469
DB 421 GTGATCTCTCCACCCAGAGGAGCAGCTCTTTCTCGCGCTGGGGCGCGCTCTCCATC 480
QY 470 CACAGTCTGGAGGCTCAGAGCCTGCGCTGCTGCAGCCCTTCTGTACCCAGGACGCGG 529
DB 481 CACAGTCTGGAGGCTCAGAGCCTGCGCTGCTGCAGCCCTTCTGTACCCAGGACGCGG 540
QY 530 GATAGGCTTTTCAGGCGCAGCCGACGAGCCTGAGACCTGCTGCTGCGGACCCCTCA 589
DB 541 GATAGGCTTTTCAGGCGCAGCCGACGAGCCTGGAAGCTGCTGCTGCGGACCCCTCA 600
QY 590 GGCCTGGTGGTGGAGAACGAAGACCGGAGACCGCGCTGGTTTCCAGCGCCCTACCTG 649

DB 601 GGCTGGTGGCTGGTGGAGAACGAAGACCGGACAGCCGCTGGTTTCCAGCGCCCTACCTG 660
QY 650 GAGGAGCGCGCCCGCGGCCAAGCGCGGAGGAGGCCCGCTCCCTAGGAGAGCGGTCCC 709
DB 661 GAGGAGCGCGCCCGCGGCCAAGCGCGGAGGAGGCCCGCTCCCTAGGAGAGCGGTCCC 720
QY 710 CAGTTCTGTGCTTCCCGCGCTACGAGAGCAGCGCGCAGATGAGCTGCTCGTGCCTCCGCG 769
DB 721 CAGTTCTGTGCTTCCCGCGCTACGAGAGCAGCGCGCAGATGAGCTGCTCGTGCCTCCGCG 780
QY 770 GGGCGCGCGTGGCGCTGTTGGAAACGTCAGACCGCGCTGGTGGCTATGCAAGTACG 827
DB 781 GGGCGCGCGTGGCGCTGTTGGAAACGTCAGACCGCGCTGGTGGCTATGCAAGTACG 838
RESULT 8
ADM02780/c
ID ADM02780 standard; cDNA; 2922 BP.
XX
AC ADM02780;
XX
DT 20-MAY-2004 (first entry)
XX
DE Human cDNA of the invention SEQ ID NO:1465.
XX
XX ss; gene; human; gene therapy; diagnostic marker; pharmaceutical.
XX
OS Homo sapiens.
XX
PN EPI347046-A1.
XX
PD 24-SEP-2003.
XX
PF 12-APR-2002; 2002EP-00008400.
XX
PR 22-MAR-2002; 2002JP-00137785.
XX
PA (REAS-) RES ASSOC BIOTECHNOLOGY.
XX
PI Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;
PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;
PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y;
XX
DR WPI; 2003-723558/69.
DR P-PSDB; ADM05223.
XX
PT New polynucleotides and polypeptides are useful in gene therapy, for
PT developing a diagnostic marker or medicines for regulating their
PT expression and activity, or as a target of gene therapy.
XX
PS Claim 1; SEQ ID NO 1465; 305pp; English.
XX
XX The invention relates to a novel human polynucleotide and the encoded
XX polypeptide. A polynucleotide of the invention may have a use in gene
XX therapy. An oligonucleotide of the invention ADM06202-ADM06773 is useful
XX as a primer for synthesizing the polynucleotide or as a probe for
XX detecting the polynucleotide. The polynucleotides ADM01316-ADM03758 are
XX useful in gene therapy, for developing a diagnostic marker or medicines
XX for regulating their expression and activity, or as a target of gene
XX therapy. The proteins ADM03759-ADM06201 encoded by the polynucleotides
XX are useful as pharmaceutical agents. The present sequence represents a
XX cDNA sequence of the invention.
SQ Sequence 2922 BP; 492 A; 954 C; 980 G; 496 T; 0 U; 0 Other;
Query Match 30.2%; Score 342.4; DB 11; Length 2922;
Best Local Similarity 84.9%; Pred. No. 4.8e-58;
Matches 428; Conservative 0; Mismatches 1; Indels 75; Gaps 1;
QY 704 GGTCCCCAGTTCTGTGTTTCCCGCGCTACGAGAGCAGCCGCGAGATGAGCTGTCCGTG 763
DB 679 GGTCCCCAGTTCTGTGTTTCCCGCGCTACGAGAGCAGCCGCGAGATGAGCTGTCCGTG 620

Qy	764	CCCGCGGGGCGCGCGTGTGGAAACGTGCGCGCTGGTGGCTAT-----	818
Db	619	CCCGCGGGGCGCGCGTGTGGAAACGTGCGCGCTGGTGGCTATGCAAG	560
Qy	819	-----	818
Db	559	TACCGGGAGCGGCGTGGGCAGGCGCAGGCTAGCCGAGGCGGCGGCCCTTAACCAAC	500
Qy	819	-----GCAGGTACGGCAGCCGGCGGCGCTACTCCCGCGGTGCTGCTGGCGCG	868
Db	499	CCCGCCCTCGCAGGTACGGCGACCGGCGGCGCTACTCCCGCGGTGCTGCTGGCGCG	440
Qy	869	GAAGGCTGGGCGCTCTCTGAGCGGA	928
Db	439	GAAGGCTGGGCGCTCTCTGAGCGGACGGGGTTCGGTGAGGAGACGACCCGCGGCT	380
Qy	929	GAGGCCCGGGCTTCCTGTGAACCTCCAGGCCACCGCCCTCCCGCCACCGTGGCCAC	988
Db	379	GAGGCCCGGGCTTCCTGTGAACCTCCAGGCCACCGCCCTCCCGCCACCGTGGCCAC	320
Qy	989	CGACTTCGCGGGCGCCATCCAGAGCGGTGTGCAACCGTCAACAGGCGCCTGGAG	1048
Db	319	CGACTTCGCGGGCGCCATCCAGAGCGGTGTGTCACCGTCAACAGGCGCCTGGAG	260
Qy	1049	CGCGCCACCGCGCCAGGCGCGCCCTCGAGGGTGCGTGACTGTGCGCACCCACCG	1108
Db	259	CGCGCCACCGCGCCAGGCGCGCCCTCGAGGGTGCGTGACTGTGCGCACCCACCG	200
Qy	1109	ACGGAGCAGTGAGCGCGAGGATCC	1132
Db	199	ACGGAGCAGTGAGCGCGAGGATCC	176

RESULT 9

AA53491	
ID	AA53491 standard; DNA; 114955 BP.
XX	
AC	AA53491;
XX	
DT	05-JUL-1999 (first entry)
XX	
DE	Human adenosine A1 receptor antisense oligonucleotide fragment.
XX	
DE	Antisense oligonucleotide; multiple target; antisense treatment;
XX	impaired respiration; inflammation; lung disease;
KW	pulmonary vasoconstriction; inflammation; allergic rhinitis;
KW	acute asthma; allergy; asthma; impeded respiration;
KW	respiratory distress syndrome; pain; cystic fibrosis;
KW	pulmonary hypertension; pulmonary vasoconstriction; emphysema;
KW	chronic obstructive pulmonary disease; leukemia; lymphoma; carcinoma;
KW	colon cancer; breast cancer; lung cancer; pancreatic cancer;
KW	hepatocellular carcinoma; kidney cancer; melanoma; hepatic metastasis;
KW	prostate cancer; ss.
XX	
OS	Synthetic.
XX	
PN	WO9913886-A1.
XX	
PD	25-MAR-1999.
XX	
PF	17-SEP-1998; 98WO-US019419.
XX	
PR	17-SEP-1997; 97US-0059160P.
PR	09-JUN-1998; 98US-00093972.
XX	
XX	(UYEC-) UNIV EAST CAROLINA.
PA	
XX	
PI	Nyce JW;
XX	
WI	1999-229400/19.
DR	
PT	New antisense oligonucleotides used in treatment of, e.g. pulmonary

PT vasoconstriction.

XX Disclosure; Page 37; 120pp; English.

PS

XX

XX The specification describes antisense oligonucleotides (AA552869-X55271) and directed against at least 2 mRNAs selected from target genes, coding and non-coding regions of RNAs corresponding to target genes, gene initiation codons, genomic flanking regions, intron-exon borders, the 5'-end, the 3'-end and the juxta-section between coding and non-coding regions and all segments of RNAs encoding proteins associated with one or more diseases, conditions or mixtures. The antisense oligonucleotides may be derived from sequences AA55272-74. These multiple target oligonucleotides (specifically AA55180-271) can be used for the antisense treatment of diseases and conditions. Typical diseases and conditions are those associated with impaired respiration and inflammation, including lung diseases, pulmonary vasoconstriction, inflammation, allergic rhinitis, acute asthma, allergies, asthma, impaired respiration, respiratory distress syndrome, pain, cystic fibrosis, pulmonary hypertension, pulmonary vasoconstriction, emphysema, chronic obstructive pulmonary disease (COPD), and cancers such as leukemias, lymphomas, carcinomas e.g. colon cancer, breast cancer, lung cancer, pancreatic cancer, hepatocellular carcinoma, kidney cancer, melanoma, hepatic metastases, as well as all types of cancers which may metastasize or have metastasized to the lungs, including breast and prostate cancer

XX Sequence 114955 BP; 6071 A; 29417 C; 36712 G; 21328 T; 0 U; 21427 Other;

XX

	Query Match	6.8%; Score 77.2; DB 2; Length 114955;
	Best Local Similarity 33.4%; Pred. No. 7.9e-06;	
	Matches 296; Conservative 92; Mismatches 484; Indels 14; Gaps 4;	
Qy	251 GGCGCAGACGCGCGCTGGCGCGCTGCAGCTGTGTGGAACCTATTCTCGAGGCTG 310 : : : : : : : : : : :	
Dd	104410 GGSNNNDNCCBGBCBGGCGCGCGCGCGCGCSNNNDNCCGBGCCBGGCGCG 104469 : : : : : : : : : : :	
Qy	311 CTGCGCACTGCAGAGCGCTGGCACGAGCCCGAC---GATCACTGGCTTCTTCGCACC 366 : : : : : : : : : : :	
Dd	104470 CCGCGGCCSNNNDNCCBGBCBGGCGCGCGCGCSNNNDNCCGBGCCBGG 104529 : : : : : : : : : : :	
Qy	367 GC AAC CCTTGACTGAGCGCGCGCTGCCACCCGGCAGCCGGTGATCTCTGCCACCCC 426 : : : : : : : : : : :	
Dd	104530 CGCGCGCGCGSNNNDNCCBGBCBGGCGCGCGCGCSNNNDNCCGBGCCBGG 104589 : : : : : : : : : : :	
Qy	427 AGAGGACGCTTTCTCGCTCGCGCGCGCGCTCTCCAATCACAGCTCTGGAGGCTCA 486 : : : : : : : : : : :	
Dd	104590 GCGCGCGCGSNNNDNCCBGBCBGGCGCGCGCGCSNNNDNCCGBGCCBGGCG 104649 : : : : : : : : : : :	
Qy	487 GAGCTCGCTGCTGCAGCCCT-----TCTGTACCAAGACACGCGGATAGCCCTTT 540 : : : : : : : : : : :	
Dd	104650 GCGCGSNNNDNCCBGBCBGGCGCGCGCSNNNDNCCBGBCBGGCGCGCSNN 104709 : : : : : : : : : : :	
Qy	541 TCAGGCGCGCCAGAGAGACTGTGAGTGCTCTCTGGCGCACCCCTCAGCGTGGTGCT 600 : : : : : : : : : : :	
Dd	104710 NDNNCCCBGGCBCGGCGCSNNNDNCCBGBCBGGCGCGCSNNNDNCCBGGCC 104769 : : : : : : : : : : :	
Qy	601 GTTGGAGAACGAAGACCGGACAGCCGCTGTGTTTTACAGCGCCT--ACCTGGAGAGCG 658 : : : : : : : : : : :	
Dd	104770 BGGCGSNNNDNCCBGBCBGGCGSNNNDNCCBGBCBGGCGSGNSNNDNCCGCBG 104829 : : : : : : : : : : :	
Qy	659 GCCCGCGCCAAAGCCGGGAGGAGCCCCCTCTCTTAGGAGACGGTCCCCAGTTCTGT 718 : : : : : : : : : : :	
Dd	104830 GCBGGSNNNDNCCBGBCBGGCSNNNDNCCBGBCBGGCSNNNDNCCGBGCCSNN 104889 : : : : : : : : : : :	
Qy	719 GCTTCCCGCGCTACGAGACGCGCGCATAGCTGTCCGTGCCCGCGGGCGCGC 778 : : : : : : : : : : :	
Dd	104890 NDNNCGCBGGCBGGCGCGCGCGCGCSNNNDNNGCBGGCBCGGCGCGCGC 104949 : : : : : : : : : : :	
Qy	779 GTGCGCTGTGGAAAAGTCAGACCGCGGCTGTGGGCTATCGAGTACGGCGACCGGCG - 837 : : : : : : : : : : :	
Dd	104950 CCGCCGGCGSNNNDNCCBGBCBGGCGCGCGCGCGCGCSNNNDNBBGGCC 105009 : : : : : : : : : : :	
Qy	838 -GGCGCTA TCCCGCGCGTGTCTGTGCGCGAGAAGGCTGGGCGCTCTCTCTGAGCGGA 896 : : : : : : : : : : :	

XX dithp.
XX Homo sapiens.
XX WO2004023973-A2.
XX 25-MAR-2004.
XX 12-SEP-2003; 2003WO-US028227.
XX 12-SEP-2002; 2002US-0410259P.
XX 12-SEP-2002; 2002US-0410260P.
XX (INCYTE) INCYTE CORP.
XX Schmidt JP, Wright RJ, Bruns CM, Marjanovic MM, Shen F, Harthorne TA, Suchorolski MT, Altus CM, Pitts SJ, Elder LV, Mooney EM, Deleane AM, Panesar IS, Banville SC, Reddy TP, Stevens KA, Blanchard JL, Rioux P, Wang X, Au AP, Gerstin EH, Peralta CH, Anderson SB, Rioux P, Shen EJ, Wu MC, Stuve LL, Lagace RE, Spiro PA, Stewart EA, Wingrove J, Vitt UA, Kirtan ES, Xu Y, Kwong M, Policky JL, Hurwitz BL, Ma Y, Jackson JL, Gietzen D, Patury S, Shi X, Suarez CJ;
XX WPI; 2004-329368/30.
XX P-PSDB; ABM94256.
XX New diagnostic and therapeutic polynucleotides and polypeptides, useful in diagnosing a condition, disease or disorder associated with human molecules, e.g. autoimmune or inflammatory disorders, in gene therapy or in gene mapping.
XX Claim 1; Page; 190pp; English.
XX The invention relates to novel diagnostic and therapeutic polynucleotides selected from one of the 2722 sequences defined in the specification. A polynucleotide of the invention may have a use in gene therapy. The human diagnostic and therapeutic polynucleotides (dithp) or polypeptides may be used to diagnose a particular condition, disease or disorder associated with human molecules, e.g. cell proliferative disorders, autoimmune/inflammatory disorder, developmental disorders, endocrine disorder, neurological disorders, gastrointestinal disorders, or infections caused by virus, bacteria, fungi or parasite. The dithp molecules may also be used in genetic mapping, in identifying individuals from minute biological samples, in detecting single nucleotide polymorphisms, as molecular weight markers, and for somatic or germline gene therapy. The present sequence represents a dithp polynucleotide of the invention. Note: The sequence data for this patent is not represented in the printed specification, but was obtained in electronic format directly from WIPO at www.wipo.int/pct/en/sequences/listing.htm
XX Sequence 2196 BP; 415 A; 708 C; 672 G; 401 T; 0 U; 0 Other;
Query Match 4.9%; Score 55.6; DB 13; Length 2196;
Best Local Similarity 49.3%; Pred. No. 0.12;
Matches 145; Conservative 0; Mismatches 149; Indels 0; Gaps 0;
QY 325 GCGCGTGGCCAGCGCCGCGGATCACTGCTTTCTTCGACCCCAACCCCTGACCTGGA 384
DB 308 GCCACCGCCGTCGCGCGCGCCGCGCCATGAGCGCGCGCGCGCGGTCGCGGTCGCT 367
QY 385 GCGCGGCTGCCACCGCGCGCGGCGGATCTCTGCGCCACCCAGAGAGAGCGCTCTTTC 444
DB 368 TCGCTGCTGCTGCTGCGCGCGCTTGGCTGCTGCGCGCGGAGTGACGCGATGCTCT 427
QY 445 TCGCGTGGCGCGCGCTCTTCATCATCAGTCTGAGAGGTCTAGAGCGCTGCGCTGCA 504
DB 428 CTGAGGCGCTGTGTGCGGACGACCGGATGCGGCACTCATCAGAGGACTGCTGCT 487
QY 505 GCCTTCTGTACCGAGACCGCGGATAGGCTTTTCAGGCGCGAGGAGCGCT 564
DB 488 GCATATGTACGAATCCAAAGATGCAGGATGGTGCAGGAGCATGCTGCCACAGCCA 547

QY 565 GGACGTGCTGCTGCGGACCCCTCAGGCTGGTGGTGGAGAACGAGACCG 618
DB 548 GCTGGAGGAGCTGCACTGTGCGCACGCGCATCAGCTGGCCACGAGCAGGACCG 601
RESULT 15
ABD33013
ID ABD33013 standard; cDNA; 2215 BP.
XX
AC ABD33013;
XX
DT 18-NOV-2004 (first entry)
XX
DE Human cancer-associated cDNA HR21-032.4.
XX
KW Human; ss; cancer-associated protein; gene; cytostatic; cancer; leukemia; lymphoma; CAP.
XX
OS Homo sapiens.
XX
PN WO2004074320-A2.
XX
PD 02-SEP-2004.
XX
PF 17-FEB-2004; 2004WO-US004730.
XX
PR 14-FEB-2003; 2003US-00367094.
PR 14-MAR-2003; 2003US-00388838.
PR 15-APR-2003; 2003US-00417375.
PR 13-JUN-2003; 2003US-00461862.
PR 15-SEP-2003; 2003US-00663431.
PR 15-DEC-2003; 2003US-00737318.
XX
PA (SAGR-) SAGRES DISCOVERY INC.
XX
PI Morris DW, Morris DW, Malandro MS;
XX
XX WPI; 2004-652914/63.
DR P-PSDB; ABO84711.
XX
XX New isolated cancer-associated polynucleotides and polypeptides useful for diagnosing, preventing or treating cancers, especially lymphoma and leukemia, or in screening for agents that modulate cancer.
XX claim 1; seqid 789; 310pp; English.
XX The invention relates to an isolated nucleic acid comprising at least 10 contiguous nucleotides of any of the 233 polynucleotide sequences given in the specification, or its complement. The nucleic acids encode cancer-associated proteins. Also included are an expression vector comprising the isolated nucleic acid cited above, a host cell comprising the above recombinant nucleic acid or expression vector, a microarray for detecting a cancer-associated (CA) nucleic acid comprising at least one probe comprising at least 10 contiguous nucleotides of any of the above-mentioned nucleotide sequences, an isolated polypeptide (encoded within an open reading frame of a CA sequence selected from any of the 95 polynucleotide sequences as mentioned in the specification, or its complement), an isolated antibody, (or its antigen binding fragment) that binds to the above polypeptide, a hybridoma that produces the above monoclonal antibody, a pharmaceutical composition comprising the above antibody and a pharmaceutical excipient, a kit for detecting cancer cells (comprising the antibody cited above, methods for diagnosing cancer or for detecting the presence or absence of cancer cells in an individual, a method for inhibiting growth of cancer cells in an individual, a method for delivering a therapeutic agent to cancer cells in an individual, an electronic library comprising the above polynucleotide or polypeptide (or their fragments), methods of screening for anticancer activity or for a bioactive agent capable of modulating the activity of a CA protein (CAP), methods for detecting cancer associated with expression of a polypeptide in a test cell sample, a method for treating cancers and a method for inhibiting the expression of CA gene in a cell. The composition and methods are useful for detecting, diagnosing, preventing and treating cancers, especially lymphoma and

CC leukaemia. These may also be used in screening for agents that modulate
CC cancer. The present sequence is a human CAP cDNA sequence. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences
XX

SQ Sequence 2215 BP; 420 A; 689 C; 665 G; 441 T; 0 U; 0 Other;

Query Match 4.9%; Score 55.6; DB 13; Length 2215;
Best Local Similarity 49.3%; Pred. No. 0.12;
Matches 145; Conservative 0; Mismatches 149; Indels 0; Gaps 0;

QY	325	GCGGTGGCAGCGAGCCGAGCATCACTGGCTTCTTCGACGCGCAACCCCTGGACCTGGA	384
Db	83	GCCACCCGCGTGCCTCCCGCCGCCCATGGAGCGCGCGCGCTCGCGCGGTCCT	142
QY	385	GCCCGCGCTGCCACCCGCGAGCCGCGGTGATCCTGTGCCACCCCGAGGAGCAGCCTCTTTC	444
Db	143	TCCGCTGTCTGTCTCGGCGCTTTCGCTGTGGCGCGGAGTGGACGCGGATGTCCT	202
QY	445	TGCGCTCGGCGCGCGCTCTCCATCCACAGTCTGGAGGCTCAGAGCCTGCGGTGCTGCA	504
Db	203	CCTGGAGCCTGCTGTGCGGACGGACACCGGATGGCCACTCATCAGAGGACTGCTCGCT	262
QY	505	GCCCTTCTGTACCCAGGACACGCGGATAGGCTTTTCAGGCGCAGGCCAGGAGCCT	564
Db	263	GCCATATGCTAGGATCCAAAGATGCAAGATGTCAGGAGCAGTGTGTCACAGCCA	322
QY	565	GGACGTGTCTGCGGCAACCCCTCAGGCTGGTGGCTGGTGGAGAACGAAGACCG	618
Db	323	GCTGGAGGAGTGCACTGTGCCACGGCATCAGCCTGGCCCAACGAGCAGGACCG	376

Search completed: May 30, 2005, 19:22:07
Job time : 978 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 30, 2005, 15:53:13 ; Search time 8111 Seconds
(without alignments)
6762.582 Million cells

Title: US-10-621-113-3

Perfect score:

Sequence: 1 agccatggcaggccccgat.....agcagtgagcgcgaggatcc 1132

Scoring table: IDENTITY NUC

scoring cases: IDENTITY_NOC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Post-processing: Minimum Match 0%
Maximum Match 100%

Maximum Macch 100%
Listing first 45 summaries

Database : GenEmbl: *

1: qb ba:*

```
1: gb_da:
2: gb_htg:*
```

3: gb_in: *

4: gb_om: *

5: gb_ov:†

6: gb_pat: *

7: gb_ph:*

8: gb_pl: *

9: gb_pr:*

10: gb_ro: *

```
11: gb_sts:*
```

```
12: gb_ey:*
13: gb_ey:*
```

```
13: gb_un:*
14: gb_un:*
```

```
14: gb_v1:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
1	1132	100.0	1132	9	AF532984	AF532984	Homo sapi
2	1132	100.0	1446	9	AB097667	AB097667	Homo sapi
3	1116	98.6	1116	9	AF539796	AF539796	Homo sapi
4	1116	98.6	1129	9	AF532983	AF532983	Homo sapi
5	1116	98.6	1587	9	BC015917	BC015917	Homo sapi
6	1107	97.8	1147	9	AF532985	AF532985	Homo sapi
7	1100	97.2	1113	9	AF525768	AF525768	Homo sapi
8	1091	96.4	1144	9	AY191359	AY191359	Homo sapi
9	526	46.5	1260	10	BC019525	BC019525	Mus muscu
10	519.6	45.9	1050	10	AF539797	AF539797	Mus muscu
11	415.4	36.7	759	10	AB105906	AB105906	Cavia por
12	410.4	36.3	412	6	CQ730806	CQ730806	Sequence
13	378.8	33.5	720	10	AF399754	AF399754	Mus muscu
14	344	30.4	80662	9	AC005606	AC005606	Homo sapi
15	342.4	30.2	2922	6	AX834341	AX834341	Sequence
16	342.4	30.2	2922	9	AC096738	AC096738	Homo sapi
17	291.6	25.8	176290	2	AC148564	AC148564	Papio anu
18	291.6	25.8	185895	2	AC150732	AC150732	Papio anu
19	201	17.8	4948	9	AY212269	AY212269	Homo sapi

LATERNVARSPTITGFPAPQDLLEPALPGSRVILPTPEOPLSRAAGRLSHSLEA
QSLRCLQPDCTQDTPEDRPOACQESLDVLLRHPSGWLVENEDRQTAFFPAPYLEEA
APGQREGGSLGSSGPFQFCASRAYESSRADELSPAGARVRVLETSDRGHMLCRYGD
RAGLLPALLRPEGLALUSGTGFRGGDDPAGEARGFPEPSQATAPPVTPTRPSPGA
IQSRCTVTTRALLERRRQRPGRGCVDSVPHPTTEQ"
17. .370
misc_feature /note="Region: PX domain"
misc_feature 485. .661
misc_feature /note="Region: SH3 domain"
707. .874
/note="Region: SH3 domain"
ORIGIN
Query Match 100.0%; Score 1132; DB 9; Length 1132;
Best Local Similarity 100.0%; Pred. No. 7.3e-177; Indels 0; Gaps 0;
Matches 1132; Conservative 0; Mismatches 0;
QY 1 AGCCATGCGAGGCCCGCCGATACCCAGTTTTCAGTGCAGAGGGGAGCCCTTGGTGCAGATCAA 60
DB 1 AGCCATGCGAGGCCCGCCGATACCCAGTTTTCAGTGCAGAGGGGAGCCCTTGGTGCAGATCAA 60
QY 61 GAGGCTCAAACGTTTGGCTTCTCTGTGGCTGTGTGAGAGCGGACGACACTTCTGTGCG 120
DB 61 GAGGCTCAAACGTTTGGCTTCTCTGTGGCTGTGTGAGAGCGGACGACACTTCTGTGCG 120
QY 121 CAGGAGTTGGGACGAATTCAGGACGCTCAAGAGACCTCAAGGAGACCTTCCCGGTGGA 180
DB 121 CAGGAGTTGGGACGAATTCAGGACGCTCAAGAGACCTCAAGGAGACCTTCCCGGTGGA 180
QY 181 GCGGGCTGTCTGCGGAGATCTGACCGGCTTCTCCAAAGTCTTCGATGACCACTGTT 240
DB 181 GCGGGCTGTCTGCGGAGATCTGACCGGCTTCTCCAAAGTCTTCGATGACCACTGTT 240
QY 241 GGGAGCGTGGGGCGCAAGAGCGGCGCTGCGCGCTGTGAGCTGTGGAAACCTATTC 300
DB 241 GGGAGCGTGGGGCGCAAGAGCGGCGCTGCGCGCTGTGAGCTGTGGAAACCTATTC 300
QY 301 TCGGAGGCTGTGGCGACTTCAGAGCGGCTGTGCGAGCGGCGGACGATCACTGGCTTCTT 360
DB 301 TCGGAGGCTGTGGCGACTTCAGAGCGGCTGTGCGAGCGGCGGACGATCACTGGCTTCTT 360
QY 361 CGCAGCGCAACCCCTGGACCTGGAGCGCGCTGCGACCCCGGACCGCGGTGATCCTGCC 420
DB 361 CGCAGCGCAACCCCTGGACCTGGAGCGCGCTGCGACCCCGGACCGCGGTGATCCTGCC 420
QY 421 CACCCAGAGAGAGAGCTTTTCTGCGCTGTGCGGGCGGCTTCCATCCACAGTCTGGA 480
DB 421 CACCCAGAGAGAGAGCTTTTCTGCGCTGTGCGGGCGGCTTCCATCCACAGTCTGGA 480
QY 481 GGCTCAGAGCTGTGCTGCTGCGAGCCCTTCTGTACCCAGAGACGCGGGATAGGCTTTT 540
DB 481 GGCTCAGAGCTGTGCTGCTGCGAGCCCTTCTGTACCCAGAGACGCGGGATAGGCTTTT 540
QY 541 TCAGGCGAGGCGCCAGGAGAGCTTGAGAGCTGTGCTGCGGCGACCCCTCAGGCTGTGCT 600
DB 541 TCAGGCGAGGCGCCAGGAGAGCTTGAGAGCTGTGCTGCTGCGGCGACCCCTCAGGCTGTGCT 600
QY 601 GGTGAGAACGAAGACCGGCGAGACCGCTGTGTTTCCAGCGCCCTACCTTGGAGGAGCGCGC 660
DB 601 GGTGAGAACGAAGACCGGCGAGACCGCTGTGTTTCCAGCGCCCTACCTTGGAGGAGCGCGC 660
QY 661 CCGGGGCAAGCGCGGAGGAGGCGGCTTCCCTAGGAGAGCGGCTTCCAGGTTCTGTGC 720
DB 661 CCGGGGCAAGCGCGGAGGAGGCGGCTTCCCTAGGAGAGCGGCTTCCAGGTTCTGTGC 720
QY 721 TTCCCGCGCTACGAGAGCAGCGCGCAGAGAGCTGTGCTGCGGCGGCGGCGCGCGT 780
DB 721 TTCCCGCGCTACGAGAGCAGCGCGCAGAGAGCTGTGCTGCGGCGGCGGCGCGCGT 780
QY 781 GCGCGTGTGGAAACGTCAGACCGCGGCTGTGCTGTATGCAAGGTACGCGCGACCGGCGGG 840
DB 781 GCGCGTGTGGAAACGTCAGACCGCGGCTGTGCTGTATGCAAGGTACGCGCGACCGGCGGG

QY 841 CCTACTCCCCCGCTGTCTGCGCGCGAAGGCTTGGCGCTCTCTCTGAGCGGACGGG 900
DB 841 CCTACTCCCCCGCTGTCTGCGCGCGAAGGCTTGGCGCTCTCTCTGAGCGGACGGG 900
QY 901 GTTCCGTGGAGAGACGACGCCGCGGGTGTAGCCCGGGGCTTCCCTGAACCTTCCAGGC 960
DB 901 GTTCCGTGGAGAGACGACGCCGCGGGTGTAGCCCGGGGCTTCCCTGAACCTTCCAGGC 960
QY 961 CACGCGCCCTCCCCCGACCGGTGCCCGACCTTCCGCGGCGGCGCATCCAGAGCGGCTG 1020
DB 961 CACGCGCCCTCCCCCGACCGGTGCCCGACCTTCCGCGGCGGCGCATCCAGAGCGGCTG 1020
QY 1021 CTGCACCGTCAACGCGAGGGCCCTGAGAGCGGCGCCACGCGGCGCGCCCTCCGAGG 1080
DB 1021 CTGCACCGTCAACGCGAGGGCCCTGAGAGCGGCGCCACGCGGCGCGCCCTCCGAGG 1080
QY 1081 GTGCGTGGACTTCTGTGCGGACCCACGACGAGCAGTGAGCGGAGGATCC 1132
DB 1081 GTGCGTGGACTTCTGTGCGGACCCACGACGAGCAGTGAGCGGAGGATCC 1132
RESULT 2
AB097667 Homo sapiens NOX01 mRNA for NADPH oxidase organizer 1, complete
LOCUS cds. 1446 bp mRNA linear PRI 03-JUL-2003
ACCESSION AB097667
VERSION AB097667.1 GI:30844229
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Takeya,R., Ueno,N., Kami,K., Taura,M., Kohjima,M., Izaki,T.,
Nunoi,H. and Sumimoto,H.
TITLE Novel human homologues of p47phox and p67phox participate in
activation of superoxide-producing NADPH oxidases
J. Biol. Chem. 278 (27), 25234-25246 (2003)
JOURNAL
MEDLINE 22716264
PUBMED 12716910
REFERENCE 2 (bases 1 to 1446)
AUTHORS Takeya,R., Taura,M. and Sumimoto,H.
TITLE Direct Submission
JOURNAL Submitted (05-DEC-2002) Ryo Takeya, Mediacal Institute of
Bioregulation; 3-1-1, Maidashi, Fukuoka, Higashi-ku 812-8582, Japan
(E-mail:takeya@bioreg.kyushu-u.ac.jp. Tel:81-92-642-6811,
Fax:81-92-642-6813)
FEATURES
source Location/Qualifiers
1. .1446
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
gene 1. .1446
/gene="NOX01"
181. .1296
/gene="NOX01"
/note="p41nox, a homologue of p47phox"
/codon_start=1
/product="NADPH oxidase organizer 1"
/protein_id="BAC76711.1"
/db_xref="GI:30844230"
translation="MAGPRPVSVQGAALVQIKELQTFAPSVRMSDSDTFFVRSWDE
FRQIKTKETFFVEAGLLRRSRKLVLPALDAPLLGRVTSRGLARLQLLETYSRRL
LATARVARSPTITGFPAPQDLLEPALPGSRVILPTPEOPLSRAAGRLSHSLEA
QSLRCLQPDCTQDTPEDRPOACQESLDVLLRHPSGWLVENEDRQTAFFPAPYLEEA
APGQREGGSLGSSGPFQFCASRAYESSRADELSPAGARVRVLETSDRGHMLCRYGD
RAGLLPALLRPEGLALUSGTGFRGGDDPAGEARGFPEPSQATAPPVTPTRPSPGA
IQSRCTVTTRALLERRRQRPGRGCVDSVPHPTTEQ"
ORIGIN
Query Match 100.0%; Score 1132; DB 9; Length 1446;
Best Local Similarity 100.0%; Pred. No. 6.8e-177;

Matches 1132; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY	1	AGCCATGGCAGGCCCCCGATACCCAGCTTTTCAGTCAAGGGCAGCGCCTCGTGCAGATCAA	60						
DB	177	AGCCATGGCAGGCCCCCGATACCCAGCTTTTCAGTCAAGGGCAGCGCCTCGTGCAGATCAA	236						
QY	61	GAGGCTCCAAACGTTTGGCTTCTCTGTGGCTGGTGCAGCGGAGGACACCTTTCGTGGG	120						
DB	237	GAGGCTCCAAACGTTTGGCTTCTCTGTGGCTGGTGCAGCGGAGGACACCTTTCGTGGG	296						
QY	121	CAGGAGTGTGGAGCAATTCAGGCAGCTCAAGAGACCTCAAGGAGACCTTCCCGGTGA	180						
DB	297	CAGGAGTGTGGAGCAATTCAGGCAGCTCAAGAGACCTCAAGGAGACCTTCCCGGTGA	356						
QY	181	GGCGGCTCTGCTGCGAGATCTGACCGCTTCTCCAAAGCTTCTCGATGACCACTGT	240						
DB	357	GGCGGCTCTGCTGCGAGATCTGACCGCTTCTCCAAAGCTTCTCGATGACCACTGT	416						
QY	241	GGACCGCTGGGCGCACAGCGCGGCTGGCGGCTGAGCTGTGGAAACCTATTTC	300						
DB	417	GGACCGCTGGGCGCACAGCGCGGCTGGCGGCTGAGCTGTGGAAACCTATTTC	476						
QY	301	TCGGAGCTGTGGGACTGCGAGCGCTGGCAGCGGCGGACCGACCATCTGGCTTCT	360						
DB	477	TCGGAGCTGTGGGACTGCGAGCGCTGGCAGCGGCGGACCGACCATCTGGCTTCT	536						
QY	361	CGCACCGCAACCTCTGGACTGAGCGCGGCTGCCACCGGCGGCGGTGATCTGCG	420						
DB	537	CGCACCGCAACCTCTGGACTGAGCGCGGCTGCCACCGGCGGCGGTGATCTGCG	596						
QY	421	CACCCAGAGGAGAGCCTTCTCGCGCTGGGCGCGCTTCCATCCAGATCTGA	480						
DB	597	CACCCAGAGGAGAGCCTTCTCGCGCTGGGCGCGCTTCCATCCAGATCTGA	656						
QY	481	GGCTCAGAGCTCGCTGCTGAGCGCTTCTGTACCCAGGACACCGGGATAGGCTTT	540						
DB	657	GGCTCAGAGCTCGCTGCTGAGCGCTTCTGTACCCAGGACACCGGGATAGGCTTT	716						
QY	541	TCAGGCGCAGGCGCCAGGAGAGCTGACCTGTGCTGCGGCGACCCCTCAGGCTGGGCT	600						
DB	717	TCAGGCGCAGGCGCCAGGAGAGCTGACCTGTGCTGCGGCGACCCCTCAGGCTGGGCT	776						
QY	601	GTGGAGAACGAGACCGGAGACCGCTGTTTCCAGCGCCCTACTCGAGGAGCGGC	660						
DB	777	GTGGAGAACGAGACCGGAGACCGCTGTTTCCAGCGCCCTACTCGAGGAGCGGC	836						
QY	661	CCGCGCCAAAGCGCGGAGGAGCCGCTCCCTAGGAGAGCGGCTCCAGTTCGTGC	720						
DB	837	CCGCGCCAAAGCGCGGAGGAGCCGCTCCCTAGGAGAGCGGCTCCAGTTCGTGC	896						
QY	721	TTCCCGCGCTACGAGAGAGCGCGCAGATGAGCTGCTCGCGCGGCGGCGCGCT	780						
DB	897	TTCCCGCGCTACGAGAGAGCGCGCAGATGAGCTGCTCGCGCGGCGGCGCGCT	956						
QY	781	GGCGGTGTTGAAACGTCAGACCGCGCTGGTGGCTATCAGGTACGGGACCGGCGGG	840						
DB	957	GGCGGTGTTGAAACGTCAGACCGCGCTGGTGGCTATCAGGTACGGGACCGGCGGG	1016						
QY	841	CCTACTCCCGCGGTGCTGCTGCGCGGAGAGGGCTGGGGCTCTCTCAGCGGAGCGGG	900						
DB	1017	CCTACTCCCGCGGTGCTGCTGCGCGGAGAGGGCTGGGGCTCTCTCAGCGGAGCGGG	1076						
QY	901	GTTTCGTGGAGAGAGACCGCGGCTGAGGCGCGGGCTTCCCTGAACCTCCAGGC	960						
DB	1077	GTTTCGTGGAGAGAGACCGCGGCTGAGGCGCGGGCTTCCCTGAACCTCCAGGC	1136						
QY	961	CACCGGCTTCCCGCCACCGCTGCCACCGACCTTCGCGGGCGGCGCATCCAGAGCGCTG	1020						
DB	1137	CACCGGCTTCCCGCCACCGCTGCCACCGACCTTCGCGGGCGGCGCATCCAGAGCGCTG	1196						
QY	1021	CTGCAACGTCACAGCGGCGCTGAGCGGCGGCCCAAGCGGCGGCGGCGCTCGAGG	1080						
DB	1197	CTGCAACGTCACAGCGGCGCTGAGCGGCGGCCCAAGCGGCGGCGGCGCTCGAGG	1256						

QY	1081	GTGCGTGGACTCTGTGCGCACCCACGAGCAGTGCAGCGGAGATCC	1132																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																										
----	------	---	------	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

Db 241 CGCGTGGGCGGCACGAGCCGCGGCTGGCGGCTGCAGCTGTTGGAACCTATTCTCGG 300
Qy 305 AGGCTGCTGGGACTGTCAGAGGCGGTGSCACGGAGCCCGACGATCACTGGCTTCTTCGCA 364
Db 301 AGGCTGCTGGGACTGTCAGAGGCGGTGSCACGGAGCCCGACGATCACTGGCTTCTTCGCA 360
Qy 365 CGCAAAACCCCTGGACCTGGAGCCCGCTGTCACACCGGACGCGGGTGCATCTGCCACAC 424
Db 361 CGCAAAACCCCTGGACCTGGAGCCCGCTGTCACACCGGACGCGGGTGCATCTGCCACAC 420
Qy 425 CCAGAGGACGACCTCTTTCTCGGCTGCGGCGGCTCTCCATCCACAGTCTGGAGCT 484
Db 421 CCAGAGGACGACCTCTTTCTCGGCTGCGGCGGCTCTCCATCCACAGTCTGGAGCT 480
Qy 485 CAGAGCTGCGTGCCTGAGGCTTCTGTATCCAGGACGACGCGGATAGGCTTTTCAG 544
Db 481 CAGAGCTGCGTGCCTGAGGCTTCTGTATCCAGGACGACGCGGATAGGCTTTTCAG 540
Qy 545 GCGCAGGCCCCAGGAGCTGTGACGTGCTGTGCGGCACCCCTCAGGCTGGTGGCTGGTG 604
Db 541 GCGCAGGCCCCAGGAGCTGTGACGTGCTGTGCGGCACCCCTCAGGCTGGTGGCTGGTG 600
Qy 605 GAGAACGAAGACCCGACAGCCCTGTTTCAGCGCCCTACCTGAGGAGGCGGCCCCG 664
Db 601 GAGAACGAAGACCCGACAGCCCTGTTTCAGCGCCCTACCTGAGGAGGCGGCCCCG 660
Qy 665 GGCNAAGCCGCGGAGGAGGCCCTCCCTAGGAGCAGCGGTCCCGAGTCTGTGCTTCC 724
Db 661 GGCNAAGCCGCGGAGGAGGCCCTCCCTAGGAGCAGCGGTCCCGAGTCTGTGCTTCC 720
Qy 725 CGCGCTACGAGAGCAGCCGCGCAGATGATGCTGCTGCGGCGGCGCGCTGCGC 784
Db 721 CGCGCTACGAGAGCAGCCGCGCAGATGATGCTGCTGCGGCGGCGCGCTGCGC 780
Qy 785 GTGTTGAAACGTCAGACCGCGGCTGTGGCTATGCAAGTACGCGACCGGCGCGGCTA 844
Db 781 GTGTTGAAACGTCAGACCGCGGCTGTGGCTATGCAAGTACGCGACCGGCGCGGCTA 840
Qy 845 CTCGCCGCTGTCTGCTGCGGCGGAAGGCTGGGCGCTTCTCTGAGCGGACGCGGTT 904
Db 841 CTCGCCGCTGTCTGCTGCGGCGGAAGGCTGGGCGCTTCTCTGAGCGGACGCGGTT 900
Qy 905 CGTGAGGAGACGACCCGCGGCTGAGGCGCGGCTTCCCTGAACCTCCACGCGCAC 964
Db 901 CGTGAGGAGACGACCCGCGGCTGAGGCGCGGCTTCCCTGAACCTCCACGCGCAC 960
Qy 965 GCCCTCCCGCCACCGTGCACCCGACCTTTCGCGGCGGCTATCCAGAGCGCTGTGTC 1024
Db 961 GCCCTCCCGCCACCGTGCACCCGACCTTTCGCGGCGGCTATCCAGAGCGCTGTGTC 1020
Qy 1025 ACCGTACACGAGGCGCTGGAGGCGGCGGCTTTCGCGGCGGCTATCCAGAGGCTGC 1084
Db 1021 ACCGTACACGAGGCGCTGGAGGCGGCGGCTTTCGCGGCGGCTATCCAGAGGCTGC 1080
Qy 1085 GTGGACTCTGTGCGGCGACCCGACGAGGAGTGA 1120
Db 1081 GTGGACTCTGTGCGGCGACCCGACGAGGAGTGA 1116

RESULT 4
AF532983
LOCUS AF532983 1129 bp mRNA linear PRI 03-JAN-2003
DEFINITION Homo sapiens regulatory protein NOX01-alpha mRNA, complete cds;
alternatively spliced.
ACCESSION AF532983
VERSION AF532983.1 GI:22532410
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1129)
AUTHORS Lambeth,J.D. and Cheng, G.

NOX01-alpha: A homolog of p47phox
Unpublished
2 (bases 1 to 1129)
Lambeth,J.D. and Cheng,G.
Direct Submission
Submitted (29-JUL-2002) Pathology, Emory University Medical School,
615 Michael St., Atlanta, GA 30322, USA
FEATURES
source
1..1129
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="16"
/map="16p13.3"
5..1117
/note="similar to p47phox; alternatively spliced"
/codon_start=1
/product="regulatory protein NOX01-alpha"
/protein_id="AAM97925.1"
/db_xref="GI:22532411"
/translation="MAGPRYPVSVQGAALVQIKELQTFAPSVRMSDGSDFVRSWDE
FRQKLUKTFPVEAGLRRSDRVLPKLLDAPLGRVGRSRLARLQLLTYSRLL
ATATVARSPTITGFFAPLDLEPALPPGSRVILPTPEQPLSRAAGRLSIHSLAQ
SLRCLQFCTQDTRDFQAQESLDVLRHPSGWLWVENEDQTAFPPAYLEEA
PGQREGPSILGSGPQFCASRAYESSRADELSPAGARVRVLETSRGMWLCRYGDR
AGLLPAVLLRPEGLALLSGTFRGGDDPAGEARGFPEPSQATAPPTVTPTRPSGAI
QSRCTVTTRALERPRRQGRGCVDSVPHPTTEQ"
17..367
/note="Region: PX domain"
482..658
/note="Region: SH3 domain"
704..880
/note="Region: SH3 domain"
ORIGIN
Query Match 98.6%; Score 1116; DB 9; Length 1129;
Best Local Similarity 99.7%; Pred. No. 3.1e-174;
Matches 1129; Conservative 0; Mismatches 0; Indels 3; Gaps 1;
Qy 1 AGCCATGGCAGGCCCGATACCCAGTTTCAGTCAAGGGGCGAGCCCTGGTGCAGATCAA 60
Db 1 AGCCATGGCAGGCCCGGATACCCAGTTTCAGTCAAGGGGCGAGCCCTGGTGCAGATCAA 60
Qy 61 GAGGCTCCAAACGTTTGCCTTCTCTGTCGCTGTGTCAGCGCAGCAGCACCTTCGTGCG 120
Db 61 GAGGCTCCAAACGTTTGCCTTCTCTGTCGCTGTGTCAGCGCAGCAGCACCTTCGTGCG 120
Qy 121 CAGGAGTTGGGACGAATTCAGGCGAGCTCAAGAAGACCTTCAAGGAGACCTTCCCGGTGA 180
Db 121 CAGGAGTTGGGACGAATTCAGGCGAGCTCAAGGAGACCTTCCCGGTGA 177
Qy 181 GCGGGGCTGTGCGGAGATCTGACCGGTTCTCCAAAGCTTCTCGATGACACCTGTT 240
Db 178 GCGGGGCTGTGCGGAGATCTGACCGGTTCTCCAAAGCTTCTCGATGACACCTGTT 237
Qy 241 GGGACGCTGGGGCGCAGGCGCGCTGCGCGCTGTCAGCTGTGGAAACCTATTTC 300
Db 238 GGGACGCTGGGGCGCAGGCGCGCTGCGCGCTGTCAGCTGTGGAAACCTATTTC 297
Qy 301 TCGGAGGCTGTGCGGAGCTGCAGAGCGGTGGGACGGAGCCCGACGATCACTGGCTTCTT 360
Db 298 TCGGAGGCTGTGCGGAGCTGCAGAGCGGTGGGACGGAGCCCGACGATCACTGGCTTCTT 357
Qy 361 CGCACCGAACCCCTGGACCTGGAGCCGCGCTGCCACCGGAGCGGGGTATCCTGCC 420
Db 358 CGCACCGCAACCCCTGGACCTGGAGCCCGCGCTGCCACCGGAGCGGGGTATCCTGCC 417
Qy 421 CACCCAGAGGAGCAGCTCTTTCTCGCGCTGCGGCGCGCTCTCCATCCACAGTCTGGA 480
Db 418 CACCCAGAGGAGCAGCTCTTTCTCGCGCTGCGGCGCGCTCTCCATCCACAGTCTGGA 477
Qy 481 GGCTCAGAGCCTGCGCTGCCTGACGCCCTTGTATCCAGGACACGCGGGATAGGCCCTTT 540

Db	478	GGCTCAGAGCCTCGGCTGCTGCTGAGCCCTTCTGTATCCAGGACACGCGGATAGGCTTT	537
Qy	541	TCAGGGCCAGCCAGGAGAGCTGACGCTGCTGCTGCGGCACCCCTCAGGCTGTGGCT	600
Db	538	TCAGGCGAGCCAGGAGAGCTGAGCTGCTGCTGCGGCACCCCTCAGGCTGTGGCT	597
Qy	601	GGTGGAGAACGAGACCGGAGACCGGCTGTGGTTTTCAGCGCCCTACTGTGAGGAGCGGC	660
Db	598	GGTGGAGAACGAGACCGGAGACCGGCTGTGGTTTTCAGCGCCCTACTGTGAGGAGCGGC	657
Qy	661	CCCGGCGCAAGGCGGAGGAGCGGCTCCCTAGGAGAGCAGCGTCCCGAGTCTGTGC	720
Db	658	CCCGGCGCAAGGCGGAGGAGCGGCTCCCTAGGAGAGCAGCGTCCCGAGTCTGTGC	717
Qy	721	TTCCCGCGCTACGAGAGCAGCGCGCAGATGAGCTGTCCGTGCCCGGGGCGGCGCT	780
Db	718	TTCCCGCGCTACGAGAGCAGCGCGCAGATGAGCTGTCCGTGCCCGGGGCGGCGCT	777
Qy	781	GCGCGTGTGGAAAGCTCAGACCGCGGCTGTGGCTGTATGAGTACGGGACCGCGGGG	840
Db	778	GCGCGTGTGGAAAGCTCAGACCGCGGCTGTGGCTGTATGAGTACGGGACCGCGGGG	837
Qy	841	CCTACTCCCGCGGCTGCTGCGCGCGGAGGCTGGGCGCTCTCTGAGCGGAGCGG	900
Db	838	CCTACTCCCGCGGCTGCTGCGCGCGGAGGCTGGGCGCTCTCTGAGCGGAGCGG	897
Qy	901	GTTCCTGTGAGGAGAGACACCGCGCGGTGAGGCGCGGGCTTCCCTGAACCTCCAGGC	960
Db	898	GTTCCTGTGAGGAGAGACACCGCGCGGTGAGGCGCGGGCTTCCCTGAACCTCCAGGC	957
Qy	961	CACCGCGCTTCCCGCACCGTGCACCGACCTTCGCGCGGCGCATCCAGAGCGCGTG	1020
Db	958	CACCGCGCTTCCCGCACCGTGCACCGACCTTCGCGCGGCGCATCCAGAGCGCGTG	1017
Qy	1021	CTGCACCTCAGCAGGAGCGCTGAGCGGCGCCACCGCGCGAGCGCGCCCTCAGG	1080
Db	1018	CTGCACCTCAGCAGGAGCGCTGAGCGGCGCCACCGCGCGAGCGCGCCCTCAGG	1077
Qy	1081	GTGCGTGGACTCTGTGCGCGCACCCACGAGCAGTGTGAGCGGAGGATCC	1132
Db	1078	GTGCGTGGACTCTGTGCGCGCACCCACGAGCAGTGTGAGCGGAGGATCC	1129
RESULT 5			
BC015917			
LOCUS			
DEFINITION			
Homo sapiens NADPH oxidase organizer 1, transcript variant a, mRNA			
BC015917			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
Homo sapiens (human)			
ORGANISM			
REFERENCE			
AUTHORS			
Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, P.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.P., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Ustin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullany, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S.G., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Rodchenko, Y., Bouffard, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalls, D.E.,			

Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.	
Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences	
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)	
12477932	
2 (bases 1 to 1587)	
Strausberg,R.	
Direct Submission	
Submitted (15-OCT-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 1A03, Bethesda, MD 20892-2590, USA	
NIH-MGC Project URL: http://mgc.nci.nih.gov	
On Aug 25, 2003 this sequence version replaced gi:16198472.	
Contact: MGC help desk	
Email: cgapbs-r@mail.nih.gov	
Tissue Procurement: ATCC	
cDNA Library Preparation: Rubin Laboratory	
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)	
DNA Sequencing by: Genome Sequence Centre, BC Cancer Agency, Vancouver, BC, Canada	
info@bcgsc.bc.ca	
Steve Jones, Sarah Barber, Mabel Brown-John, Yaron Butterfield, Andy Chan, Steve S. Chand, William Chow, Alison Cloutier, Ruth Featherstone, Malachi Griffith, Obi Griffith, Ran Guin, Nancy Liao, Kim MacDonald, Amara Masson, Mike R. Mayo, Josh Moran, Ryan Morin, Teika Olson, Diana Palmquist, Anca Petrescu, Anna Liisa Prabhu, parvaneh Saesedi, JR Santos, Angelique Schnerch, Ursula Skalska, Duane Smailus, Jeff Stott, Miranda Tsai, George Yang, Jacque Schein, Asim Siddiqui, Rob Holt, Marco Marra.	
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov	
Series: IRAL plate: 29 Row: p Column: 12	
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 21389404.	
Location/Qualifiers	
1..1587	
/organism="Homo sapiens"	
/mol_type="mRNA"	
/db_xref="taxon:9606"	
/clone="MGC:20258 IMAGE:4661469"	
/tissue_type="Colon, adenocarcinoma"	
/clone_lib="NIH MGC 15"	
/lab_host="DH10B-R"	
/note="Vector: pOTB7"	
1..1587	
/gene="NOXO1"	
/note="synonyms: P41NOX, P41NOXA, P41NOXB, P41NOXC, MGC20258"	
/db_xref="LocusID:124056"	
261..1373	
/gene="NOXO1"	
/codon_start=1	
/product="NADPH oxidase organizer 1, isoform a"	
/protein_id="AAH15917.1"	
/db_xref="GI:16198473"	
/db_xref="LocusID:124056"	
/translation="MAGPRYFVSOGAALVQIKLQTFATFVSVDGSDTFVRRSWDE FQLKLTETPEVAGLRSDRVLPKLLDAPLGLRGVTRGLARQLLETYSRRLL ATAERVARSPTITGFPQPLDLEPALPGSRVILPTPEEQPLSRAAGRLSHLEAQ SIRCLOPFCTDTRDRPFOAQAESLDVLRHPSGWLWVNEEDRQTAMFPAPYLEEA PQGRRGFSLSGGSPQFCASRAYESSRADELSPAGARVRVLETSDRGWMLCRYGDR AGLLPVLLRPEGLGALLSGTGFRCGDDPAGEARGFPEPSPQATPPTVTPRPSGAI QSRCTCTRRALERPRRQGRPCRCVDSVPHPTTEQ"	
ORIGIN	
Query Match 98.6%; Score 1116; DB 9; Length 1587;	
Best Local Similarity 99.7%; Pred. No. 2.9e-174;	
Matches 1129; Conservative 0; Mismatches 0; Indels 3; Gaps 1;	
1 AGCCATGCGAGCCCGCCGATACCCAGTTTCAGTGCAGAGCGGCGCCCTCGTCAGATCAA 60	

Db 257 AGCATGGCAGGCCCCCGATACCAGTTTTCAGTGCAGGGGCAGCCCTGGTGCAGATCAA 316

Qy 61 GAGGCTCAAACGTTTGCTTCTCTGTGGCTGGTTCAGACGGCAGCGACACCTTTCGTGG 120

Db 317 GAGGCTCAAACGTTTGCTTCTCTGTGGCTGGTTCAGACGGCAGCGACACCTTTCGTGG 376

Qy 121 CAGGAGTTGGGACGAATTCAGGCAGCTCAAGAAGACCTCAAGGAGACCTTCCCGGTGGA 180

Db 377 CAGGAGTTGGGACGAATTCAGGCAGCTC---AAGACCTCAAGGAGACCTTCCCGGTGGA 433

Qy 181 GCGGGCTGCTGCGGAGACTGACCGGTTTCTCCAAAGCTTTCGATGACACCACTGTT 240

Db 434 GCGGGCTGCTGCGGAGACTGACCGGTTTCTCCAAAGCTTTCGATGACCACTGTT 493

Qy 241 GCGACGCTGGGGCCACGAGCGCGGCTGGCGGCTCGAGCTGTGGAGTGGAAACCTATTC 300

Db 494 GCGACGCTGGGGCCACGAGCGCGGCTGGCGGCTCGAGCTGTGGAAACCTATTC 553

Qy 301 TCGGAGGCTGCTGGCGACTGACAGCGGTGGACGGAGCCGACGATCACTGGCTTCTT 360

Db 554 TCGGAGGCTGCTGGCGACTGACAGCGGTGGACGGAGCCGACGATCACTGGCTTCTT 613

Qy 361 CGCACCGCAACCCCTGGACCTGGAGCCCGGTGTCACCCCGCAGCGGGTGATCCTGCC 420

Db 614 CGCACCGCAACCCCTGGACCTGGAGCCCGGTGTCACCCCGCAGCGGGTGATCCTGCC 673

Qy 421 CACCCACAGAGAGAGCTTCTTCTGGCGCTGGCGGCGGCTTCTCCATCCACAGTCTGGA 480

Db 674 CACCCACAGAGAGAGCTTCTTCTGGCGCTGGCGGCGGCTTCTCCATCCACAGTCTGGA 733

Qy 481 GGCTCAGAGAGCTGCGCTGCGAGCCCTTCTGTACCCAGAGACACGGGGATAGSCCTTT 540

Db 734 GGCTCAGAGAGCTGCGCTGCGAGCCCTTCTGTACCCAGAGACACGGGGATAGSCCTTT 793

Qy 541 TCAGCGCAGGCCACAGGAGAGCTTGAGAGTGTGCTGGCGGACCCCTCAGGCTGGTGCT 600

Db 794 TCAGCGCAGGCCACAGGAGAGCTTGAGAGTGTGCTGGCGGACCCCTCAGGCTGGTGCT 853

Qy 601 GGTGAGAAACGAAGACCGGAGACCGCTTGTTCAGCGGCTTACCTTGAGAGAGCGGCG 660

Db 854 GGTGAGAAACGAAGACCGGAGACCGCTTGTTCAGCGGCTTACCTTGAGAGAGCGGCG 913

Qy 661 CCCGGGCAAGCGCGGAGGAGGCGCTGCTCCTAGGAGAGAGCGGTCCCCAGTTCGTGTC 720

Db 914 CCCGGGCAAGCGCGGAGGAGGCGCTGCTCCTAGGAGAGAGCGGTCCCCAGTTCGTGTC 973

Qy 721 TTCCGCGCCTACGAGAGAGCGCGAGAGTGTCCGTGCCCGCGGGGCGCGGT 780

Db 974 TTCCGCGCCTACGAGAGAGCGCGAGAGTGTCCGTGCCCGCGGGGCGCGGT 1033

Qy 781 GCGGTGTTGGAACGTGAGACCGCGGCTGTGGCTATGACAGTACGGCGACCGGGCGGG 840

Db 1034 GCGGTGTTGGAACGTGAGACCGCGGCTGTGGCTATGACAGTACGGCGACCGGGCGGG 1093

Qy 841 CCTACTCCCGCGGTGCTGCGGCGGAGAGGCTGGCGCTCTCTGAGCGGAGCGG 900

Db 1094 CCTACTCCCGCGGTGCTGCGGCGGAGAGGCTGGCGCTCTCTGAGCGGAGCGG 1153

Qy 901 GTTTCGTGGAGAGACGACCGCGGGGTGAGCGCCCGGGGCTTCCCTGAAACCTTCCAGGC 960

Db 1154 GTTTCGTGGAGAGACGACCGCGGGGTGAGCGCCCGGGGCTTCCCTGAAACCTTCCAGGC 1213

Qy 961 CACGCGCCTTCCCGCAGCGGTGCCACCCGACCTTTCGCGCGGCGGCATCCAGAGCGGTG 1020

Db 1214 CACGCGCCTTCCCGCAGCGGTGCCACCCGACCTTTCGCGCGGCGGCATCCAGAGCGGTG 1273

Qy 1021 CTGACCTCTCACGACAGGCGCTGAGCGGCGCCACGGCGCCAGGCGCCGCTCGAGG 1080

Db 1274 CTGACCTCTCACGACAGGCGCTGAGCGGCGCCACGGCGCCAGGCGCCGCTCGAGG 1333

Qy 1081 GTGCGTGAGACTGTGTGCGGCAACCCACGACGAGAGAGTGTGAGCGCGAGGATCC 1132

Db 1334 GTGCGTGAGACTGTGTGCGGCAACCCACGACGAGAGAGTGTGAGCGCGAGGATCC 1385

RESULT 6

AF532985

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

1. 1147

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/chromosome="16"

/map="16p13.3"

5. 1135

/note="similar to p47phox; alternatively spliced"

/codon_start=1

/product="regulatory protein NOXO1-gamma"

/protein_id="AA097927.1"

/db_xref="GI:22532415"

/translation="MAGPRYPVSVQGAALVOIKRLQTFAFSVRWSDGSDTFFVRKSWDE

PROLKTLETFFVEAGLRSDRLVPLKLGQSLDAPLLGRVGRTRGLARLOLLET

YSRLLATASRVARSPTITGFPAPQPLDLSPALPPGSRVILPTPEQPLSRAAGRLSI

HSLSAQRLCQPFCTQDRPQAQQLSDVLLRHPGSMWLVENEDRTANFPAP

YLBEAAPGQREGFSLGSSGQPCASRAYESSRDELSPAGARVVRVLETSDRGMWL

CTYDRAGLPLVALLRPEGLGALSGTFRGGDDPAGEARGFPSPQATAPPTVPTPR

PSPAIQSRQCTVTRALERRRPRRQGRPRGCVDSVPHPPTTEQ"

misc_feature

17. 385

/note="Region: PX domain"

misc_feature

500. 676

/note="Region: SH3 domain"

misc_feature

722. 889

/note="Region: SH3 domain"

ORIGIN

Query Match

Best Local Similarity

Matches 1132; Conservative

97.8%; Score 1107; DB 9; Length 1147;

98.7%; Pred. No. 9.5e-173;

Mismatches 0; Indels 15; Gaps 1;

Qy 1 AGCCATGCGAGGCCCGCATACCCAGTTTCAGTGCAGGGGCAGCCCTGGTGCGATCAA 60

Db 1 AGCCATGCGAGGCCCGCATACCCAGTTTCAGTGCAGGGGCAGCCCTGGTGCGATCAA 60

Qy 61 GAGGCTCCAAACGTTTGCTTCTCTGTGGCTGGTTCAGACGGCAGCGACACCTTCGTGG 120

Db 61 GAGGCTCCAAACGTTTGCTTCTCTGTGGCTGGTTCAGACGGCAGCGACACCTTCGTGG 120

Qy 121 CAGGAGTTGGAGCGAATTCAGGCGAGCTCAAGAAGACCTTCAAGAGACCTTCCCGGTGGA 180

Db 121 CAGGAGTTGGAGCGAATTCAGGCGAGCTCAAGAAGACCTTCAAGAGACCTTCCCGGTGGA 180

Qy 181 GCGCGGCTGCTGGGAGACTGACCGGTTCTCCCAAGCTTCTC----- 226

Db 181 GCGCGGCTGCTGGGAGACTGACCGGTTCTCCCAAGCTTCTCGGTGAGGCCAGGCT 240

Qy 227 -GATGCACCACTGTTGGGAGCGGTGGGGCGCACGAGCGCGGCTTGGCGCGCTTGCAGCT 285

AF532985 1147 bp mRNA linear PRI 03-JAN-2003

Homo sapiens regulatory protein NOXO1-gamma mRNA, complete cds;

alternatively spliced.

AF532985

AF532985.1 GI:22532414

Homo sapiens (human)

Homo sapiens

Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 1147)

Lambeth, J.D. and Cheng, G.

NOXO1-gamma: A homolog of p47phox

Unpublished

2 (bases 1 to 1147)

Lambeth, J.D. and Cheng, G.

Direct Submission

Submitted (29-JUL-2002) Pathology, Emory University Medical School,

615 Michael St., Atlanta, GA 30322, USA

Location/Qualifiers

1. 1147

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/chromosome="16"

/map="16p13.3"

5. 1135

/note="similar to p47phox; alternatively spliced"

/codon_start=1

/product="regulatory protein NOXO1-gamma"

/protein_id="AA097927.1"

/db_xref="GI:22532415"

/translation="MAGPRYPVSVQGAALVOIKRLQTFAFSVRWSDGSDTFFVRKSWDE

PROLKTLETFFVEAGLRSDRLVPLKLGQSLDAPLLGRVGRTRGLARLOLLET

YSRLLATASRVARSPTITGFPAPQPLDLSPALPPGSRVILPTPEQPLSRAAGRLSI

HSLSAQRLCQPFCTQDRPQAQQLSDVLLRHPGSMWLVENEDRTANFPAP

YLBEAAPGQREGFSLGSSGQPCASRAYESSRDELSPAGARVVRVLETSDRGMWL

CTYDRAGLPLVALLRPEGLGALSGTFRGGDDPAGEARGFPSPQATAPPTVPTPR

PSPAIQSRQCTVTRALERRRPRRQGRPRGCVDSVPHPPTTEQ"

misc_feature

17. 385

/note="Region: PX domain"

misc_feature

500. 676

/note="Region: SH3 domain"

misc_feature

722. 889

/note="Region: SH3 domain"

ORIGIN

Query Match

Best Local Similarity

Matches 1132; Conservative

97.8%; Score 1107; DB 9; Length 1147;

98.7%; Pred. No. 9.5e-173;

Mismatches 0; Indels 15; Gaps 1;

Qy 1 AGCCATGCGAGGCCCGCATACCCAGTTTCAGTGCAGGGGCAGCCCTGGTGCGATCAA 60

Db 1 AGCCATGCGAGGCCCGCATACCCAGTTTCAGTGCAGGGGCAGCCCTGGTGCGATCAA 60

Qy 61 GAGGCTCCAAACGTTTGCTTCTCTGTGGCTGGTTCAGACGGCAGCGACACCTTCGTGG 120

Db 61 GAGGCTCCAAACGTTTGCTTCTCTGTGGCTGGTTCAGACGGCAGCGACACCTTCGTGG 120

Qy 121 CAGGAGTTGGAGCGAATTCAGGCGAGCTCAAGAAGACCTTCAAGAGACCTTCCCGGTGGA 180

Db 121 CAGGAGTTGGAGCGAATTCAGGCGAGCTCAAGAAGACCTTCAAGAGACCTTCCCGGTGGA 180

Qy 181 GCGCGGCTGCTGGGAGACTGACCGGTTCTCCCAAGCTTCTC----- 226

Db 181 GCGCGGCTGCTGGGAGACTGACCGGTTCTCCCAAGCTTCTCGGTGAGGCCAGGCT 240

Qy 227 -GATGCACCACTGTTGGGAGCGGTGGGGCGCACGAGCGCGGCTTGGCGCGCTTGCAGCT 285

QY 425 CCAGAGGAGCAGCTCTTTCTCGCGCTCGGGCCGCTCTCCATCCAGTCTGAGGCT 484
| | | | |
Db 418 CCAGAGGAGCAGCTCTTTCTCGCGCTCGGGCCGCTCTCCATCCAGTCTGAGGCT 477
| | | | |
QY 485 CAGAGCTCGCGTGCCTGCGAGCCCTTCTGTATCCAGGACACGCGGGATAGGCTTTTCAG 544
| | | | |
Db 478 CAGAGCTCGCGTGCCTGCGAGCCCTTCTGTATCCAGGACACGCGGGATAGGCTTTTCAG 537
| | | | |
QY 545 GCGAGGCGCAGGAGAGCTTGACGTGCTGTGCGGACACCCCTCAGGCTGTGCTGTG 604
| | | | |
Db 538 GCGAGGCGCAGGAGAGCTTGACGTGCTGTGCGGACACCCCTCAGGCTGTGCTGTG 597
| | | | |
QY 605 GAGACGAAAGCCGACAGACGCTGTTTCTCAGGCGCTTACTCTGAGGAGCGGCCCG 664
| | | | |
Db 598 GAGACGAAAGCCGACAGACGCTTGGTTTTCAGGCGCTTACTCTGAGGAGCGGCCCG 657
| | | | |
QY 665 GSCCAAGCGCGGAGGAGCGCCGCTCCCTAGGAGACAGCGGTCCCAGTTCCTGTCTTCC 724
| | | | |
Db 658 GSCCAAGCGCGGAGGAGCGCCGCTCCCTAGGAGACAGCGGTCCCAGTTCCTGTCTTCC 717
| | | | |
QY 725 CGCGCTACGAGAGCAGCGCGCAGATGAGTGTCCGTGCGCGGGGGCGGCGTGC 784
| | | | |
Db 718 CGCGCTACGAGAGCAGCGCGCAGATGAGTGTCCGTGCGCGGGGGCGGCGTGC 777
| | | | |
QY 785 GTGTTGGAACGTACAGACCGGCGTGTGGCTATGACGTACGCGACCGGGCGGCCTA 844
| | | | |
Db 778 GTGTTGGAACGTACAGACCGGCGTGTGGCTATGACGTACGCGACCGGGCGGCCTA 837
| | | | |
QY 845 CTCCTCCGCGGTGCTGCTGCGCGGAGGCGTGGCGCTCTCTGAGCGGACGCGGTTTC 904
| | | | |
Db 838 CTCCTCCGCGGTGCTGCTGCGCGCGGAGGCGTGGCGCTCTCTGAGCGGACGCGGTTTC 897
| | | | |
QY 905 COTGAGAGAGCAGACCCGCGGCTGAGGCGCGGCTTCCCTGAACCTCCAGGCGCAC 964
| | | | |
Db 898 COTGAGAGAGCAGACCCGCGGCTGAGGCGCGGCTTCCCTGAACCTCCAGGCGCAC 957
| | | | |
QY 965 GCCCTCCCGCCACCGTGCACCGACCTTCGCGCGGGCGCATCCAGAGCGCTGCTGC 1024
| | | | |
Db 958 GCCCTCCCGCCACCGTGCACCGACCTTCGCGCGGGCGCATCCAGAGCGCTGCTGC 1017
| | | | |
QY 1025 ACCGTACACGAGGCGCTTGAGCGGCGCCACGCGCGCCAGGCGCGCTTGAGGCTGC 1084
| | | | |
Db 1018 ACCGTACACGAGGCGCTTGAGCGGCGCCACGCGCGCCAGGCGCGCTTGAGGCTGC 1077
| | | | |
QY 1085 GTGGACTGTGCGCCACCCACGACGAGCAGTGTA 1120
| | | | |
Db 1078 GTGGACTGTGCGCCACCCACGACGAGCAGTGTA 1113
| | | | |

RESULT 8
AY191359 1144 bp mRNA linear PRI 15-FEB-2003
LOCUS
DEFINITION Homo sapiens regulatory protein NOX01-delta (NOX01) mRNA, complete cds; alternatively spliced.
ACCESSION AY191359
VERSION AY191359.1 GI:28394745
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1144)
Cheng, G. and Lambeth, J.D.
TITLE Regulatory protein NOX01-delta
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1144)
Cheng, G. and Lambeth, J.D.
TITLE Direct Submission
JOURNAL Submitted (05-DEC-2002) Pathology, Emory University School of Medicine, 615 Michael Street, Atlanta, GA 30322, USA
FEATURES
1..1144
source
/organism="Homo sapiens"

/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="16"
/map="16p13.3"
1..1144
/gene="NOX01"
5..1132
/gene="NOX01"
/note="similar to p47phox; alternatively spliced"
/codon_start=1
/product="regulatory protein NOX01-delta"
/protein_id="AA038665.1"
/db_xref="GI:28394746"
/translation="MAGPRYPVSVOGAALVQIKELQTFAPSVRMSDGSDFVRSWDE
FRQKTKETFPVEAGLLRRSDRVLPKLLQASLDAPLGRVGTSGRLARQLLEY
SRLLIATAERVARSPITTFAPQPLEPALPFGSRVILPTPEEQPLSRAAGLSIH
SLEAQSLRCLQPFCTQDRDRPFOAQESLDVLLRHPSGMWLVENEDROTATWPPAPY
LEEAAPCGREGGSPGLSGSPQFCASRAYSSRADELSSVAGARVRVLETSDRGMWLC
RYGDRAGLLPAVLLRPGLGALLSGTGRGDDDPAGEARGFPPEPSQATAPPTVPTRP
SPGAIQSRSCCTVTRALLRRPRQGRPRGCVDSVPHPHTTEQ"

ORIGIN

Query Match 96.4%; Score 1091; DB 9; Length 1144;
Best Local Similarity 98.4%; Pred. No. 4.1e-170;
Matches 1129; Conservative 0; Mismatches 0; Indels 18; Gaps 2;
QY 1 AGCCATGGCAGGCCCGGATACCCAGTTTCAGTCGACGAGGCGACGCTTGTGCAGATCAA 60
| | | | |
Db 1 AGCCATGGCAGGCCCGGATACCCAGTTTCAGTCGACGAGGCGACGCTTGTGCAGATCAA 60
| | | | |
QY 61 GAGGCTCCAAACGTTTGCCTTCTGTGCGCTGTGTCAGTCGACGAGCGACACCTTCGTGCG 120
| | | | |
Db 61 GAGGCTCCAAACGTTTGCCTTCTGTGCGCTGTGTCAGTCGAGCGAGCGACACCTTCGTGCG 120
| | | | |
QY 121 CAGGAGTTGGGACGAATTCAGGCGAGCTCAAGAAGACCTCAAGAGAGCTTTCCCGGTGGA 180
| | | | |
Db 121 CAGGAGTTGGGACGAATTCAGGCGAGCTC--NAGACCTCAAGGAGACCTTCCCGGTGGA 177
| | | | |
QY 181 GCGCGGCTGCTGGGAGATCTGACCGGCTTCTCCAAAGCTTCTC----- 226
| | | | |
Db 178 GCGCGGCTGCTGGGAGATCTGACCGGCTTCTCCAAAGCTTCTCGGTGAGGCGAGCCT 237
| | | | |
QY 227 -GATGCACCACTGTTGGGACGCGTGGGCGCGACGAGCGCGGCTTGGCGCGCTGCAGCT 285
| | | | |
Db 238 GGATGCACCACTGTTGGGACGCGTGGGCGCGACGAGCGCGGCTTGGCGCGCTGCAGCT 297
| | | | |
QY 286 GTTCGAAACCTATTTCTCGAGGCTGCTGGCGACTGACAGAGCGCTGGGACGAGCGCCGAC 345
| | | | |
Db 298 GTTCGAAACCTATTTCTCGAGGCTGCTGGCGACTGACAGAGCGCTGGGACGAGCGCCGAC 357
| | | | |
QY 346 GATCACTGGCTTCTTCGACCGCAACCCCTGGAGCTTGGAGCCCGGCTGCCACCGGCGAG 405
| | | | |
Db 358 GATCACTGGCTTCTTCGACCGCAACCCCTGGAGCTTGGAGCCCGGCTGCCACCGGCGAG 417
| | | | |
QY 406 CCGGCTGATCTCGCCACCCCGAGAGGAGCAGCTTCTTCTCGCGCTGCGGCGCGCTCTC 465
| | | | |
Db 418 CCGGCTGATCTCGCCACCCCGAGAGGAGCAGCTTCTTCTCGCGCTGCGGCGCGCTCTC 477
| | | | |
QY 466 CATCCACAGTCTGAGGCTCAGAGCTGCGCTGCTGAGCCCTTCTGTATCCAGGAGACAC 525
| | | | |
Db 478 CATCCACAGTCTGAGGCTCAGAGCTGCGCTGCTGAGCCCTTCTGTATCCAGGAGACAC 537
| | | | |
QY 526 GCGGGATAGGCTTTTCAGGCGCAGGCCCGAGGAGGCTTGACCTGCTGCTGCGGCGACCC 585
| | | | |
Db 538 GCGGGATAGGCTTTTCAGGCGCAGGCCCGAGGAGGCTTGAGCTGCTGCTGCGGCGACCC 597
| | | | |
QY 586 CTCAGGCTGTGGCTGTGGAGAAACGAAGACCGCGAGACCGCTTGTTCAGGCGCCCTA 645
| | | | |
Db 598 CTCAGGCTGTGGCTGTGGAGAAACGAAGACCGCGAGACCGCTTGTTCAGGCGCCCTA 657
| | | | |
QY 646 CCTCGAGAGAGCGCGCCCGGCGCCAAAGCCCGGAGAGGAGGCCGCTCCTTAGGAGGAGCGG 705
| | | | |
Db 658 CCTCGAGAGAGCGCGCCCGGCGCCAAAGCCCGGAGAGGAGGCCGCTCCTTAGGAGGAGCGG 717
| | | | |

QY	706	TCCCCAGTTCTGCTTCCCGCGCTTACGAGAGCAGCGCGCAGATGAGTGTCCGCTGCC	765
Db	718	TCCCCAGTTCTGCTTCCCGCGCTTACGAGAGCAGCGCGCAGATGAGTGTCCGCTGCC	777
QY	766	CGCGGGGGCGCGTGGCGGTGTGAAACGTAGACCGCGGCTGTGGTATGACAGGTA	825
Db	778	CGCGGGGGCGCGTGGCGGTGTGAAACGTAGACCGCGGCTGTGGTATGACAGGTA	837
QY	826	CGCGGACCGGGCGGCTACTTCCCGCGGTCTGCTGCGCGCCGGAAGGGCTGGCGGCTCT	885
Db	838	CGCGGACCGGGCGGCTACTTCCCGCGGTCTGCTGCGCGCCGGAAGGGCTGGCGGCTCT	897
QY	886	CTTGAGCGGACCGGGTTCCTGAGAGAGACGACCCCGCGGGTGAAGCCCGGGGCTTCCC	945
Db	898	CTTGAGCGGACCGGGTTCCTGAGAGAGACGACCCCGCGGGTGAAGCCCGGGGCTTCCC	957
QY	946	TGAACCTCCAGCGCACCGCCCTCCCGCCACGCGTGGCCACCGACCTTCGCGGGCGC	1005
Db	958	TGAACCTCCAGCGCACCGCCCTCCCGCCACGCGTGGCCACCGACCTTCGCGGGCGC	1017
QY	1006	CATCCAGAGCGGCTGTGACCGTTCACACGACGCGGCTTGAGCGGCGCCACGGCGCCA	1065
Db	1018	CATCCAGAGCGGCTGTGACCGTTCACACGACGCGGCTTGAGCGGCGCCACGGCGCCA	1077
QY	1066	GGCGCCCTCGAGGGTGGCGTGAAGTCTGTGCGCGCACCCCGACGAGCAGTGAGCGCG	1125
Db	1078	GGCGCCCTCGAGGGTGGCGTGAAGTCTGTGCGCGCACCCCGACGAGCAGTGAGCGCG	1137
QY	1126	AGGATCC 1132	
Db	1138	AGGATCC 1144	
RESULT 9			
BC019525			
LOCUS			
DEFINITION	Mus musculus NADPH oxidase organizer 1, mRNA (cdna clone MGC:28618 IMAGE:4220018), complete cds.		
ACCESSION	BC019525		
VERSION	BC019525.1	GI:18043495	
KEYWORDS	MGC.		
SOURCE	Mus musculus (house mouse)		
ORGANISM	Mus musculus		
REFERENCE			
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 1260)		
	Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D., Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Schettin,T.E., Brownstein,M.J., Udwin,T.B., Toshiyuki,S., Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullaly,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahney,J., Helton,E., Kettelman,M., Madan,A., Rodriguez,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakeley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butterfield,Y.S., Kravinsky,M.I., Skalska,U., Smallos,D.E., Scherch,A., Schein,J.E., Jones,S.J. and Marra,M.A.		
TITLE	Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences		
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)		
MEDLINE	22388257		
PUBMED	12477932		
REFERENCE	2 (bases 1 to 1260)		
AUTHORS	Strausberg,R.		
TITLE	Direct Submission		

JOURNAL

REMARK
COMMENT

Submitted (19-DEC-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: cgabs@mail.nih.gov

Tissue Procurement: Jeffrey E. Green, M.D.

cDNA Library Preparation: Life Technologies, Inc.

DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center

Center code: BCM-HGSC

Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>

Contact: amg@bcm.tmc.edu

Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulseged, H., Kowis, C.R., Speed, A.J., Martin, R.G., Muzny, D.M., Nanavati, A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAK Plate: 37 Row: j Column: 17

This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 26080419.

FEATURES

source

1..1260

/organism="Mus musculus"

/mol_type="mRNA"

/strain="FVB/N"

/db_xref="taxon:10090"

/clone="MGC:28618 IMAGE:4220018"

/tissue_type="Colon, normal. 5 month old male mouse."

/clone_lib="NCI CGAP_Co24"

/lab_host="DH10B"

/note="Vector: PCMV-SPORT6"

gene

1..1260

/gene="Noxol"

/note="synonyms: Snx28, P41NOX, P41NOXA, P41NOXB, P41NOXC"

/db_xref="LocusID:71893"

/db_xref="MGI:1919143"

CDS

27..1076

/gene="Noxol"

/codon_start=1

/product="Noxol protein"

/protein_id="AAH19525.1"

/db_xref="GI:18043496"

/db_xref="LocusID:71893"

/db_xref="MGI:1919143"

/translation="MASPRHPVSAHAVALQMDRLQTFAFVSCWSDNSTFVRRSWDE
PRLQKTKTPFVEAGLLRSEQVLPKLPDAPLLTRRGTGRGLVRLLDYVQAL
LATSHILRSSALHGFFVKPDLPLPGLSLVLTPTPEELSQPRGSLDTHSLBAQ
STPCQPHETLDIRDPHPTKAQEILDLIRHPSGWLVENKQDQVAFPPAPVLEVA
TQCGESGLALQSGRGQCTTQAYGSRSELSVFSGARVHVLETSDRGWMLCRYNGR
TGLLPAMSLQPEGLSLGLRPGFDSAGAKVAEDRTTTPVVPVTRPCMSAIQSRCCSI
TRRALAQEQGTRVR"

93..389

/gene="Noxol"

/note="PX; Region: PX domain. PX domains bind to phosphoinositides"

/db_xref="CDD:pfam00787"

misc_feature

576..677

/gene="Noxol"

/note="SH3; Region: Src homology 3 domains"

/db_xref="CDD:smart00326"

misc_feature

747..887

/gene="Noxol"

/note="SH3; Region: Src homology 3 domains"

/db_xref="CDD:smart00326"

misc_feature

93..389

/gene="Noxol"

/note="SH3; Region: Src homology 3 domains"

/db_xref="CDD:smart00326"

ORIGIN

Query Match 46.5%; Score 526; DB 10; Length 1260;
Best Local Similarity 69.6%; Pred. No. 4, 7e-77;
Matches 753; Conservative 0; Mismatches 305; Indels 24; Gaps 2;

Db	241	CGTCGGGGGCACTACTGTCAGAGACTGGTACGTTTGGGGTGTCTGGACACCTATGTACAG	300
Qy	305	AGGCTGTGCGGACTGCAGAGCGCGTGGCAGCGAGCGCCGACGATCACTGGCTTCTTCGCA	364
Db	301	GCATTGCTGCGCACTCAGAACACATATTGAGGAGTTTCAGCACTTCACGGCTTCTTTGTA	360
Qy	365	CCGCAACCTCTGACCTGGAGCGCGGCTGCCACCGCGCAGCGGGTGATCTTGCCCAAC	424
Db	361	CCCAACCTCTGATCTGGAGCGCCATGCTGCTCTCTGGCAGCTGTGATCTCTGCTCTACA	420
Qy	425	CCAGAGGAGAGCCTCTTCTCGGCTCGGGCGGCTCTCCATCCACAGTCTGGAGGCT	484
Db	421	CCAGAGG---AGCCCTTATCCCAACCCAGAGGAGCCTTGACATTCATAGCCTGGAGGCT	477
Qy	485	CAGAGCTGTGCTGCTGCGCTCTCTGTATCCAGGACACGCGGATAGGCTTTTCAG	544
Db	478	CAGAGCATTCCTGTGTACAGCCTTCCACACTCTTGACATNAGACAGACCTTTCCAC	537
Qy	545	GCGCAGGCCAGGAGAGCTGGAGCTGTCTGCTGGGACCCCTCAGGCTGGTGCTGGTG	604
Db	538	ACCAAGGCTCAAGAAATTCTGGACATATTACTACGACATCTTTCAGGCTGGTGGCTGGT	597
Qy	605	GAGAACGAGACGGCAGACCGCTGTGTTCCAGCGCCCTACCTGAGGAGGCGGCCCG	664
Db	598	GAGAACAGATCAGCAGGTAGCTGTGTTCCAGCTCCCTTACCTGAGGAGGTAGCAACG	657
Qy	665	GGCAAGGCGGAGGAGGCGCTCTCTAGGAGCAGCGGTCCCAAGTTCTGTGTTCC	724
Db	658	TGCCAAGGCCAGGAGTCAGGCTGGCTTTGCAAGGAAGTGGGAGGCAGTTCTGCATCT	717
Qy	725	CGGCTCTAGAGAGCGCGCAGATGAGCTCTGCTGCTCCCGGGGCGCGCTGCGC	784
Db	718	CAGGCTTACAGGCGAGTCTGCTGATGAGTATCTGCTGCTCCCTCAGGGGCACTGTCCAT	777
Qy	785	GTGTTGGAACGTCAGACCGGCTGTGCTGATGAGTACGCTACCGCAGCGGGCGGCTA	844
Db	778	GTCTGGAGACCTCAGACCGAGCTGTGCTGTGCAAGGATTAATGGCCGGAAGGCGCTA	837
Qy	845	CTCCCGCGGTGTCTGCTGCGCGGAGGCTGGGCTCTCTGAGCGGAGCGGCTTC	904
Db	838	CTCCCTGCAATGTCGCTGCAACTGAAAGGCTGGGCTGCTCTGCGCAGGCGAGGCTTC	897
Qy	905	CGTGGAGGAGACGACCGCGGCTGAGCGCGGGCTTCCCTGAACCTCCCAAGGCCACC	964
Db	898	CCAGACAGTGTGGGCGAGACAAAGTGG-----CTGAGGACAGG	936
Qy	965	GCCCTCTCCCACTGCTGCCACCCGACCTTCGCGGGCGCCATCCAGAGCGCTGCTGC	1024
Db	937	ACCATTCCTCTGTATGACCACTGCTGCTGTATGAGTGCCATCCAGAGTGTGCTGC	996
Qy	1025	ACCGTCACAGCAGGCGCTTGGAGCGCGGCCACCGCGCGCGGCGGCTCGA	1078
Db	997	TCCATTACCGGAGGCACTGGGACAGGAAGGACTCGGTTTCCCGTTGA	1050
RESULT 11			
AB105906			
LOCUS			
DEFINITION			
Cavia porcellus NOXO1 mRNA for NADPH oxidase organizer 1, partial cds.			
ACCESSION			
AB105906			
VERSION			
AB105906.1 GI:42794029			
KEYWORDS			
SOURCE			
Cavia porcellus (domestic guinea pig)			
ORGANISM			
Cavia porcellus			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.			
REFERENCE			
1			
Kawahara, T., Kuwano, Y., Teshima-Kondo, S., Takeya, R., Sumimoto, H., Kishi, K., Tsunawaki, S., Hirayama, T. and Rokutan, K.			
Role of Nicotinamide Adenine Dinucleotide Phosphate Oxidase 1 in Oxidative Burst Response to Toll-Like Receptor 5 Signaling in Large Intestinal Epithelial Cells			

JOURNAL	J. Immunol. 172 (5), 3051-3058 (2004)
PUBMED	14978110
REFERENCE	2 (bases 1 to 759)
AUTHORS	Kawahara, T. and Rokutan, K.
TITLE	Direct Submission
JOURNAL	Submitted (15-MAR-2003) Kazuhito Rokutan, University of Tokushima, Department of Nutritional Physiology, School of Medicine; 3-18-15 Kuramoto-cho, Tokushima, Tokushima 770-8503, Japan (E-mail:rokutan@nutr.med.tokushima-u.ac.jp, Tel:81-88-6333-9246, Fax:81-88-6333-7086)
FEATURES	
source	Location/Qualifiers
	1..759
	/organism="Cavia porcellus"
	/mol_type="mRNA"
	/strain="Hartley"
	/db_xref="taxon:10141"
gene	1..759
	/gene="NOXO1"
CDS	<1..>759
	/gene="NOXO1"
	/note="p41NOX"
	/codon_start=1
	/product="NADPH oxidase organizer 1"
	/protein_id="BAD11766.1"
	/db_xref="GI:42794030"
	/translation="WSDGSDTFCRSWDEFRLQKTLKENFPFVEAGLLRRSDRLPKLPDAPLSRGRTGRGLVRLRLNLTVQELLANVQVMSVPLTSFPEPRNDLESALPPSMVLLPVPEKPSRPVNSLDIHSLEIQSLCCVHPFTDQGRPFHVKAQESLDVLLRHPGWLIVENEGQQKAPFPAPYLEAVESGLMLQNSGTVFCASRAYKGSRADELSVPAGARVRVLETSRQWMLCRYCDQGLLPFVLLQP"
ORIGIN	
Query Match	36.7%; Score 415.4; DB 10; Length 759;
Best Local Similarity	73.1%; Pred. No. 9e-59;
Matches 568; Conservative	0; Mismatches 191; Indels 18; Gaps 2;
Qy	92 TGGTCAGACGGCAGCAGCACTTCTGTCGCGCAGAGTTGGACGAATTCAGGCGAGCTCAAG 151
Db	1 TGGTCGGATGGCAGTGATACCTTTGTTTCAGAAAGCTGGGATGAGTTTCAGGCGAGCTCCAG 60
Qy	152 AAGACCTTCAAGGAGACCTTCCGGTGGAGCGGGCTCTGTCGGAGATCTGACCGGCTT 211
Db	61 AAAACCTTCAAGGAGAACTTTCAGTGGAGGCGAGTCTGCTGAGGAGATCTGATCGCCTT 120
Qy	212 CTCCCAAAGCTTCTCGATGCACCATGTTGGGACGCGTGGGCGCACAGCCGGCGCTG 271
Db	121 CTCCCAAGCTTCCAGATGCACCTTGTCTTCCGTGGAGGGCGCACAGTCTGTGCTG 180
Qy	272 GCGCGCTCGAGCTGTGGAACCTATTTCTGGAGGCTCTGCGGACTGTCGAGCGCGCTG 331
Db	181 GTGCGTCTGAGGCTGTGTAACACCTATGTGCAGGAACCTGCTAGCCAATGTGAACCAAGTG 240
Qy	332 GCACGAGCCGACGATCACTGGCTTCTTCGCACCGCAACCCCTGGACCTGGAGCCGCG 391
Db	241 TCATGGAGCCCAAGTGTCTTACAAGCTTCTTTGAGCCTCGAAACAGGACCTTGGAACTCTGCA 300
Qy	392 CTGCCACCGGCGACGGGGTGATCTTGCCCAACCCCAAGAGGAGCAGCTCTTTCTCGCGCT 451
Db	301 CTGCTCCCGCAGCAGATGTGTATCTGCTCTCCAGAGAGCCCTCAC---CTCGCCCT 357
Qy	452 GCGGGCGGCTCTCCATCCACAGTCTGGAGGCTCAGAGCCTGGGCTGCTCTGCAGCCCTTC 511
Db	358 GTGAACAGCTTTGATATCCACAGCTTAGAGATTCAGAGCTCTGTTCGGTGCATCTCTTC 417
Qy	512 TGTACCCAGGACACGGGGATAGGCTTTTTCAGGCGCAGCGCCCGAGAGAGCCTTGGAGCTG 571
Db	418 CACACCCAGGACACACAGGGAAGGCGCTTCCATGTGAAGGCCCAAGAAAGCCTTGGATGTA 477
Qy	572 CTGCTGCGGACCCCTCAGGCTGTGGCTGTGGTGGAGAACGAAGACCGGACAGCCGCTGG 631
Db	478 CTGCTACGACACCCCTCAGGCTGTGGCTGTGGAGNATGAAGGCCAGCAGAGGCTTGG 537
Qy	632 TTTCCAGCGCCCTTACCTTGGAGGAGGGGGCCCCCGGGCCCAAGCGCGGAGGAGGCCCGCTCC 691

```
Db 538 TTTCCAGTCCCTATCTGGAGTGGC-----CCAGAGGTGGCTTGATG 582
QY 692 CTAGGAGCAGCGGTCCCGCTTCTGTGTTCCCGGCTTACGAGAGCAGCGCGCAGAT 751
Db 583 CTGCAAAAACAGTGGGAGTGTGTCTGTCTTCCCGAGCCTACAAGGCGAGCGTGTGAT 642
QY 752 GAGCTGTCCGTCCCGCGCGGCGCGTGGCGGTGGAAACCTCAGACCGCGGCTGG 811
Db 643 GAACGTGTAGTCCAGCGGAGCAGAGTGGGGTCTTGGAAACTTCAGACCGTGGATGG 702
QY 812 TGGCTATGAGTACGCGACCGCGGCGGCTTACTCCCGCGGTGTGTGCGGCGG 868
Db 703 TGGCTGTGACAGTACTGTGACCAAGATGCTGTGCTTCACTGCTGTGCAACCG 759

RESULT 12
LOCUS CQ730806 412 bp DNA linear PAT 03-FEB-2004
DEFINITION Sequence 16740 from Patent WO02068579.
ACCESSION CQ730806
VERSION CQ730806.1 GI:42305651
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
REFERENCE
AUTHORS Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.
TITLE Kites, such as nucleic acid arrays, comprising a majority of
humanexons or transcripts, for detecting expression and other uses
thereof
JOURNAL Patent: WO 02068579-A 16740 06-SEP-2002;
PE Corporation
LOCATION (NY) (US)
FEATURES
source
1..412
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
```

```
ORIGIN
Query Match 36.3%; Score 410.4; DB 6; Length 412;
Best Local Similarity 99.8%; Pred. No. 7, 1e-58;
Matches 411; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 705 GTCCCAAGTTCGTCTTCCCGGCTTACGAGAGCAGCGCGCAGATGATGTCGTGTC 764
Db 1 GTCCCAAGTTCGTCTTCCCGGCTTACGAGAGCAGCGCGCAGATGATGTCGTGTC 60
QY 765 CCGCGGGGCGCGGTGCGCGGTGGAAACGTACAGCGCGGCTGGTGGCTATGCGAGT 824
Db 61 CCGCGGGGCGCGGTGCGCGGTGGAAACGTACAGCGCGGCTGGTGGCTATGCGAGT 120
QY 825 ACGGCGACCGGGCGGCTTACTCCCGGCTGCTGCTGCGCGGCGGAGGGCTGGCGCTC 884
Db 121 ACGGCGACCGGGCGGCTTACTCCCGGCTGCTGCTGCGCGGCGGAGGGCTGGCGCTC 180
QY 885 TCCTGAGCGGAGCAGGGTTCGTGTGAGGAGACGACCCCGCGGCTGAGCGCGGCGCTTCC 944
Db 181 TCCTGAGCGGAGCAGGGTTCGTGTGAGGAGACGACCCCGCGGCTGAGCGCGGCGCTTCC 240
QY 945 CTGAACCTTCCCGAGGACCGGCTTCCCGGCTGCGGCTGCGGCTTCCCGGCGG 1004
Db 241 CTGAACCTTCCCGAGGACCGGCTTCCCGGCTGCGGCTGCGGCTTCCCGGCGG 300
QY 1005 CCATCCAGAGCGCTGTGTCACCGGTACACGAGGCGGCTTGGAGCGGCGGCGGCGG 1064
Db 301 CCATCCAGAGCGCTGTGTCACCGGTACACGAGGCGGCTTGGAGCGGCGGCGGCGG 360
QY 1065 AGGGCGGCGCTGAGGGTGGCGTGGAGCTCTGTGCGCGGACCCCGACGAGGAGCA 1116
Db 361 AGGGCGGCGCTGAGGGTGGCGTGGAGCTCTGTGCGCGGACCCCGACGAGGAGCA 412
```

```
RESULT 13
LOCUS AF399754 720 bp mRNA linear ROD 22-AUG-2001
DEFINITION Mus musculus SNX28 (Snx28) mRNA, complete cds.
ACCESSION AF399754
VERSION AF399754.1 GI:15277910
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE
AUTHORS Hong, W
TITLE Direct Submission
JOURNAL Submitted (16-JUL-2001) Membrane Biology Laboratory, Institute of
Mol. & Cell. Biol., 30 Medical Drive, Singapore 117609
FEATURES
source
1..720
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
1..720
/gene="Snx28"
1..720
/gene="Snx28"
/feature="member of sorting nexin family"
/codon_start=1
/product="SNX28"
/protein_id="AAK94017.1"
/db_xref="GI:15277911"
/translation="MASPRHPVSAHAVALVQMDRLQTFAPSVCHSDNSDTFVRBSWDE
FRQKTLTKTFPVEAGLLRRSEQLPALDAPLIRGHTGRGLVRLRLDITVQAL
LATSEHLRSLAHGFVPKPLDLEPMLPPSLVILTPPEPLSPQPSGLDIHSLAQ
SIPCVQPFHTLDIRDPHTKAQIBLILLRHPSSMWLVENKQQQVAFWPAPYLEEVA
TCQGESGLALQSGMSPSLI"
ORIGIN
```

```
Query Match 33.5%; Score 378.8; DB 10; Length 720;
Best Local Similarity 72.2%; Pred. No. 9, 7e-53;
Matches 507; Conservative 0; Mismatches 192; Indels 3; Gaps 1;
QY 5 ATGCAGGCCCCCATACCCAGTTTCAGTGCAGAGGCGAGCCCTGTCAGATCAAGAGG 64
Db 1 ATGCAGGCCCCAAGACACCCAGTATCAGCCCATCTGTGTAGCTTGTGCAATGGACCGA 60
QY 65 CTCGAAAGCTTTGCTCTCTGTGCGTGTGTGCAGACGCGCAGCAGACCTTCGTGCGCAGG 124
Db 61 CTCGAGACATTTGCTCTCTCGTGTGTGTGTGCAGACACAGTGCACACATTTGTGCGGAGG 120
QY 125 AGTTGGAGACGAATTCAGGAGCTCAAGAGACCCCTCAAGGAGACCTTCCCGGTGAGGCGG 184
Db 121 AGCTGGGATGATTTTCAGGAGCTCCAGAGACCCCTTAAGAAAAACCTTCCAGTGGAGGCA 180
QY 185 GGCCTGTGCGGAGATCTGACCGGCTTCTCCCAAGCTTCTCGATGACCACTGTGTTGGGA 244
Db 181 GGCCTGTGACGAGATCTGAACAAAGTTCTTCCCAAGCTTCTCGATGCTCCATTTGCTGACA 240
QY 245 CGCCTGGGCGGCACAGCGCGGCTGCGCGCTGCGAGCTGTGTGGAAACCTATTCTCGG 304
Db 241 CGTGGGGGCATCTGCTGAGGACTGGTACGTTTGGGCTGCTGGACACCTATGTACAG 300
QY 305 AGGCTGTGCGGACTGACAGAGCGGTGGCAGCGAGCCCGAGCATCACTGGGCTTCTTCGCA 364
Db 301 GCATTGTGCGCAACTCAGAACACATATTGAGGAGTTTCAGCACTTCACGGCTTCTTTGTA 360
QY 365 CCGCAACCCCTGGACCTGGAGCGCGCTGCCACCCGCGAGCGGCTGATCTCGCCACCC 424
Db 361 CCCAACTCTGGATCTGGAGCCCATGCTGCTCTCTGGAGCCTTGGTGTATCTGCTTACA 420
QY 425 CCAGAGGAGCAGCCTCTTTCTCGGCTGCGGCGCGCTCTCCATCCACAGTCTGAGGCT 484
```


misc_feature /rpt_family="LTR12"
18578..18845 2 excellent exon, frame 1"
/note="GRAIL 2 excellent exon, frame 1"
repeat_region complement(19317..19602)
/rpt_family="Alu"
misc_feature 20244..20383
/note="GRAIL 2 excellent exon, frame 1"
complement(21204..21361)
misc_feature /note="GRAIL 2 excellent exon, frame 1"
complement(21588..21747)
misc_feature /note="GRAIL 2 excellent exon, frame 0"
complement(21812..22000)
misc_feature /note="GRAIL 2 excellent exon, frame 2"
complement(22240..22347)
misc_feature /note="GRAIL 2 excellent exon, frame 0"
22427..22718
repeat_region /rpt_family="Alu"
complement(23337..23456)
misc_feature /note="GRAIL 2 excellent exon, frame 1"
repeat_region complement(25378..25789)
/rpt_family="Alu"
misc_feature 26365..26510
/note="GRAIL 2 excellent exon, frame 0"
misc_feature 26816..26959
/note="GRAIL 2 excellent exon, frame 2"
misc_feature 27525..27980
/note="GRAIL 2 excellent exon, frame 2"
28148..29029
repeat_region /rpt_family="Alu"
complement(29102..29373)
repeat_region /rpt_family="Alu"
complement(29472..29999)
repeat_region /rpt_family="Alu"
complement(30309..30626)
repeat_region /rpt_family="Alu"
32251..32538
/rpt_family="Alu"
33147..33198
/note="GRAIL 2 excellent exon, frame 0"
misc_feature 33288..33383
/note="GRAIL 2 excellent exon, frame 2"
misc_feature 33467..33514
/note="GRAIL 2 excellent exon, frame 1"
33629..33770
/note="GRAIL 2 excellent exon, frame 0"
33854..33942
/note="GRAIL 2 excellent exon, frame 0"
34019..34189
/note="GRAIL 2 excellent exon, frame 2"
34436..34529
/note="GRAIL 2 excellent exon, frame 1"
34620..34801
/note="GRAIL 2 excellent exon, frame 1"
34899..35040
/note="GRAIL 2 excellent exon, frame 2"
35147..35205
/note="GRAIL 2 excellent exon, frame 0"
35905..36052
/note="GRAIL 2 excellent exon, frame 1"
36145..36375
/note="GRAIL 2 excellent exon, frame 0"
36447..36517
/note="GRAIL 2 excellent exon, frame 2"
36605..36764
/note="GRAIL 2 excellent exon, frame 2"
36833..36891
/note="GRAIL 2 excellent exon, frame 1"
37049..37152
/note="GRAIL 2 excellent exon, frame 2"
37240..37328
/note="GRAIL 2 excellent exon, frame 2"
37404..37562
/note="GRAIL 2 excellent exon, frame 1"

misc_feature complement(38516..38633)
/note="GRAIL 2 excellent exon, frame 1"
Query Match 30.4%; Score 344; DB 9; Length 80662;
Best Local Similarity 85.1%; Pred. No. 1.4e-47;
Matches 429; Conservative 0; Mismatches 0; Indels 75; Gaps 1;
QY 704 GGTCCCGAGTTCTGCTTCCGCGCCTACGAGAGCAGCGCGCAGATGAGTGTCCGTG 763
DB 38634 GGTCCCGAGTTCTGCTTCCGCGCCTACGAGAGCAGCGCGCAGATGAGTGTCCGTG 38575
QY 764 CCCGCGGGGCGCGCTGCGCGTGTGAAACGTCAGACCCGCGTGTGGCTAT----- 818
DB 38574 CCCGCGGGGCGCGCTGCGCGTGTGAAACGTCAGACCCGCGTGTGGCTATGCGAG 38515
QY 819 ----- 818
DB 38514 TACGCGGAGCGCGTGGCGAGGCGAGGCTAGCGAGCGGGCGGCGCCCTAACACAC 38455
QY 819 -----GCAGGTAGCGGCGCGGGCGGCGCTACTCCCGCGGTGCTGTCGGCGG 868
DB 38454 CCCGCCCTCGCAGGTACGGCGACCGGGCGGCGCTACTCCCGCGGTGCTGTCGGCGG 38395
QY 869 GAAGGGCTGGCGCTCTCTCTGAGCGGACGGGGTTCCGTGAGGAGACGACCCGCGGGT 928
DB 38394 GAAGGGCTGGCGCTCTCTCTGAGCGGACGGGGTTCCGTGAGGAGACGACCCGCGGGT 38335
QY 929 GAGGCCCGGGCTTCCCTGAACCTCCGAGGCCACCGGCCCTCCCGCCACCGTGCACAC 988
DB 38334 GAGGCCCGGGCTTCCCTGAACCTCCGAGGCCACCGGCCCTCCCGCCACCGTGCACAC 38275
QY 989 CGACTTTCGCGGGCGGCATCCAGAGCGCTGTCACCGTCACAGGAGGCGCTGAG 1048
DB 38274 CGACTTTCGCGGGCGGCATCCAGAGCGCTGTCACCGTCACAGGAGGCGCTGAG 38215
QY 1049 CGGCGCCACGCGCGCGCGCTCGAGGGCGCGCTCGAGGGTGGTGACTCTGTCCCGCACCCACG 1108
DB 38214 CGGCGCCACGCGCGCGCGCTCGAGGGCGCGCTCGAGGGTGGTGACTCTGTCCCGCACCCACG 38155
QY 1109 ACGAGCAGTGAGCGCGAGGATCC 1132
DB 38154 ACGAGCAGTGAGCGCGAGGATCC 38131
RESULT 15
AX834341/c 2922 bp DNA linear PAT 15-DEC-2003
LOCUS Sequence 1465 from Patent EP1347046.
DEFINITION AX834341
ACCESSION AX834341
VERSION AX834341.1 GI:39920476
KEYWORDS Homo sapiens (human)
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS
1 Isogai,T., Sugiyama,T., Otsuki,T., Wakamatsu,A., Sato,H., Ishii,S.,
Yamamoto,J.I., Isono,Y., Hio,Y., Otsuka,K., Nagai,K., Irie,R.,
Tamechika,I., Seki,N., Yoshikawa,T., Otsuka,M., Nagahari,K. and
Masuho,Y.
TITLE Full-length cDNA sequences
JOURNAL Patent: EP 1347046-A 1465 24-SEP-2003;
Research Association for Biotechnology (JP)
FEATURES
source
1. .2922
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 30.2%; Score 342.4; DB 6; Length 2922;
Best Local Similarity 84.9%; Pred. No. 6.5e-47;
Matches 428; Conservative 0; Mismatches 1; Indels 75; Gaps 1;

Qy	704	GGTCCCAGTTCTGTGTTCCCGGCTACGAGAGCAGCCGCGAGATGAGTGTCCGTG	763
Db	679	GGTCCCAGTTCTGTGTTCCCGGCTACGAGAGCAGCCGCGAGATGAGTGTCCGTG	620
Qy	764	CCGCGGGGGCGCGGTGCGGTTGGAAACGTCAACCCGCGGCTGTGGCTAT-----	818
Db	619	CCGCGGGGGCGCGGTGCGGTTGGAAACGTCAACCCGCGGCTGTGGCTATGCAGG	560
Qy	819	-----	818
Db	559	TACGCGGAGCGGGTGGGCGAGGAGGTAGCCGAGGCGGGGCCCTAACACAC	500
Qy	819	-----GCAGGTACGCGACCGGGCGGGCTACTCCCGCGGTGTGTGCGGCCG	868
Db	499	CCGCGCCCTCGCAGGTACGCGACCGGGCGGGCTACTCCCGCGGTGTGTGCGGCCG	440
Qy	869	GAAGGGCTGGCGCTCTCTGAGCGGACGGGGTTCGTGGAGAGACGACCCGCGGGT	928
Db	439	GAAGGGCTGGCGCTCTCTGAGCGGACGGGGTTCGTGGAGAGACGACCCGCGGGT	380
Qy	929	GAGGCCCGGGCTTCCCTGAACCTCCAGGCCACCGCCCTCCCGCCACCGTCCACC	988
Db	379	GAGGCCCGGGCTTCCCTGAACCTCCAGGCCACCGCCCTCCCGCCACCGTCCACC	320
Qy	989	CGACCTTCGCGGGGCGCATCCAGAGCGCTGTGCACCGTCAACGAGGGCCCTGGAG	1048
Db	319	CGACCTTCGCGGGGCGCATCCAGAGCGCTGTGCACCGTCAACGAGGGCCCTGGAG	260
Qy	1049	CGGCGCCACGCGGCGCAGGGCCGCCCTCGAGGGTGGTGAATCTGTGCGGCACCCACG	1108
Db	259	CGGCGCCACGCGGCGCAGGGCCGCCCTCGAGGGTGGTGAATCTGTGCGGCACCCACG	200
Qy	1109	ACGAGCAGTGAAGCGAGGATCC	1132
Db	199	ACGAGCAGTGAAGCGAGGATCC	176

Search completed: May 30, 2005, 21:37:31
Job time : 8117 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 30, 2005, 18:26:34 ; Search time 5919 Seconds
(without alignments)
7279.735 Million cell updates/sec

Title: US-10-621-113-3
Perfect score: 1132
Sequence: 1 agccatggcagggcccccgcatt.....agcagtgcgagcaggatcc 1132

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*

1: gb_est1:*

2: gb_est2:*

3: gb_hic:*

4: gb_est3:*

5: gb_est4:*

6: gb_est5:*

7: gb_est6:*

8: gb_gse1:*

9: gb_gse2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	524.4	46.3	2195	3	AK088226	AK088226 Mus muscu
2	466.8	41.2	790	4	BG470266	BG470266 602533785
3	432.2	38.2	519	5	BX283317	BX283317 BX283317
C	429	37.9	559	6	CA311857	CA311857 UT-CP-FN0
5	427.4	37.8	751	1	AL042759	AL042759 DRPZP434C
6	421.2	37.2	623	2	BE793330	BE793330 601588311
7	421	37.2	473	4	BG328824	BG328824 602427426
8	381.2	33.7	1571	3	AK009605	AK009605 Mus muscu
9	376.2	33.2	904	4	BG871968	BG871968 602793072
10	347	30.7	726	4	BI409553	BI409553 602963963
11	334.8	29.6	544	2	BE751524	BE751524 203692 MA
12	326.4	28.8	773	4	BG247058	BG247058 602359714
13	318.4	28.1	918	2	BF581002	BF581002 602100116
14	317.4	28.0	602	6	CB113007	CB113007 K-EST0154
15	309.2	27.3	967	5	BQ935073	BQ935073 AGENCOURT
16	304.4	26.9	584	2	AW743151	AW743151 up63D08.Y
17	300	26.5	635	1	AA543925	AA543925 vk40C12.I
18	282.4	24.9	984	6	BY762147	BY762147 BY762147
19	281.8	24.9	541	2	AW743167	AW743167 up63d12.Y
20	274.8	24.3	542	1	AA269948	AA269948 v364g11.I
21	268	23.7	274	2	AW163511	AW163511 au95f11.Y
22	258	22.8	865	7	CO810832	CO810832 AGENCOURT
23	242.6	21.4	743	4	BI083169	BI083169 602875260
24	233.8	20.7	559	1	AV594936	AV594936 AV594936

25	232	20.5	582	5	BP346589	BP346589 BP346589
26	230	20.3	480	4	BI534701	BI534701 397823 MA
C	227	19.6	349	2	BF591178	BF591178 7h43a06.X
28	217.8	19.2	450	5	BQ265615	BQ265615 NTSC ff07
29	206.4	18.2	437	2	BE554337	BE554337 ur41G11.Y
C	189.8	16.8	645	7	CN284040	CN284040 170004705
C	185.8	16.4	245	1	AA576374	AA576374 nm75a09.B
32	185.8	16.4	288	2	AW358014	AW358014 41850 MAR
C	181.4	16.0	898	5	BQ222005	BQ222005 AGENCOURT
34	180.4	15.9	837	6	BY709278	BY709278 BY709278
C	180	15.9	303	1	AI435004	AI435004 tb76g05.X
36	178.4	15.8	658	1	AL042758	AL042758 DRPZP434C
37	173	15.3	405	1	AI019507	AI019507 ua89d10.I
38	165.8	14.6	333	1	AI173340	AI173340 uc29f05.I
39	164	14.5	572	2	BF140326	BF140326 601787835
40	157.8	13.9	612	4	BI113592	BI113592 602897862
41	154	13.6	733	4	BI415175	BI415175 602990692
42	152.2	13.4	906	5	BQ959588	BQ959588 AGENCOURT
C	149.6	13.2	306	2	AW083983	AW083983 xc36g09.X
44	148.6	13.1	427	2	BF522080	BF522080 UT-R-Y0-a
45	147.8	13.1	812	4	BI112184	BI112184 602899851

ALIGNMENTS

RESULT 1	AK088226	2195 bp	linear	HTC 03-APR-2004
LOCUS	Mus musculus	2 days neonate thymus	thymic cells cDNA, RIKEN	
DEFINITION	full-length enriched library, Clone:E430007K11			
	product:2310034C04RIK PROTEIN (SNX28) homolog [Mus musculus], full			
	insert sequence.			
ACCESSION	AK088226			
VERSION	AK088226.1	GI:26353183		
KEYWORDS	HTC; CAP trapper.			
SOURCE	Mus musculus (house mouse)			
ORGANISM	Mus musculus			
REFERENCE	1			
AUTHORS	Carninci, P. and Hayashizaki, Y.			
TITLE	High-efficiency full-length cDNA cloning			
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)			
MEDLINE	99279253			
PUBMED	10349636			
REFERENCE	2			
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Sugahara, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.			
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to			
JOURNAL	prepare full-length cDNA libraries for rapid discovery of new genes			
MEDLINE	Genome Res. 10 (10), 1617-1630 (2000)			
PUBMED	20499374			
REFERENCE	3			
AUTHORS	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsumai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanabe, M., Yoneda, Y., Ishikawa, T., Tanaka, T., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.			
TITLE	RIKEN integrated sequence analysis (RISA) system--384-format			
JOURNAL	sequencing pipeline with 384 multicapillary sequencer			
MEDLINE	Genome Res. 10 (11), 1757-1771 (2000)			
PUBMED	20530913			
REFERENCE	4			
AUTHORS	The RIKEN Genome Exploration Research Group Phase II Team and the			
TITLE	PANTOM Consortium			
JOURNAL	Functional annotation of a full-length mouse cDNA collection			
REFERENCE	Nature 409, 685-690 (2001)			

AUTHORS	The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.									
TITLE	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs									
JOURNAL	Nature 420, 563-573 (2002)									
REFERENCE	6 (bases 1 to 2195)									
AUTHORS	Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T., Katoh,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M., Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N., Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akaiura,S., Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.									
TITLE	Direct Submission									
JOURNAL	Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.jp, URL:http://genome.gsc.riken.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)									
COMMENT	cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Tissues were provided by Dr. John Todd (Dept. of Medical Genetics Wellcome Trust Centre for Molecular Mechanisms in Disease Wellcome Trust/MRC building Addenbrookes Hospital Cambridge) whose assistance we gratefully acknowledge. Please visit our web site for further details. URL:http://genome.gsc.riken.jp/ Location/Qualifiers 1. .2195 /organism="Mus musculus" /mol_type="mRNA" /strain="NOD" /db_xref="FANTOM,DB:E430007K11" /db_xref="taxon:10090" /clone="E430007K11" /cell_type="thymic cells" /tissue_type="thymus" /clone_lib="RIKEN full-length enriched mouse cDNA library" /dev_stage="2 days neonate" 974. .2023 /note="unnamed protein product; 2310034C04RIK PROTEIN (SNX28) homolog [Mus musculus] (SPTR Q9D747, evidence: FASTV, 99.6%ID, 97.4%length, match=699) putative" /codon_start=1 /protein_id="BAC40222.1" /db_xref="GI:26353184" /translation="MASPRHPVSAHAVALVQMDRLQTFAPSVCSWSDNSDTFVRRSWDE FRLQKTLKTFPVVEAGLRRSEOVLPKLPDAPLLTRRGTLRGLVRLRLDLYVOAL LATSEHLIRSSALHGFVPKPLDLEPMLPGSLVILPTPEPLSQPSGLDIHSLAQ SIPQVPHLLDIRDPFHTRKAEILDLIRHPSGWLIVENKQOVAFPPAPYLEVA TCQGESGLAQGGQRFCTTQAVEGSRSDLSVPSGARVHVLFTSDRWLRCYNGR TGLPAMLSQPEGLSLGLRFPFDSAGADKVAEDRTTPPVVPRPCMSAIQSRCCSI TRRALGQSGTRVRP" 2172. .2177 /note="putative" 2195 /note="putative"									
FEATURES	source									
CDS										
polya_signal	2172. .2177									
polya_site	2195									
ORIGIN	Query Match 46.3%; Score 524.4; DB 3; Length 2195; Best Local Similarity 69.5%; Pred. No. 2.8e-98;									


```

Qy 1081 GT 1082
Db 2026 AT 2027

RESULT 2
BG470266 790 bp mRNA linear EST 21-MAR-2001
LOCUS 602533785F1 NIH_MGC_15 Homo sapiens cDNA clone IMAGE:4661469 5',
DEFINITION mRNA sequence.
ACCESSION BG470266
VERSION BG470266.1 GI:13402541
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 790)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strauberg, Ph.D.
Email: csapbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: NIH Intramural Sequencing Center
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLC1460 row: i column: 22
High quality sequence stop: 780.
FEATURES
    source
        1..790
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="IMAGE:4661469"
            /tissue_type="adenocarcinoma cell line"
            /lab_host="DH10B (phage-resistant)"
            /clone_lib="NIH_MGC_15"
            /notes="Organ: colon; Vector: pOTB7; Site 1: XhoI; Site 2:
            EcoRI; cDNA made by oligo-dt priming. Directionally
            cloned into EcoRI/XhoI sites using the following 5',
            adaptor: GGACGAG(G). Size-selected >500bp for average
            insert size 1.8kb. Library constructed by Ling Hong in
            the laboratory of Gerald M. Rubin (University of
            California, Berkeley) using ZAP-cDNA synthesis kit
            (Stratagene) and Superscript II RT (Life Technologies)"
ORIGIN
    Query Match 41.2%; Score 466.8; DB 4; Length 790;
    Best Local Similarity 98.3%; Pred. No. 2.3e-86;
    Matches 514; Conservative 0; Mismatches 2; Indels 6; Gaps 4;

Qy 1 AGCCATGGAGGCCCCCGATACCCAGTTTCAGTCAAGGGGAGCCCTGGTGAGATCAA 60
    |||||
Db 269 AGCCATGGAGGCCCCCGATACCCAGTTTCAGTCAAGGGGAGCCCTGGTGAGATCAA 328
    |||||

Qy 61 GAGGCTCCAAACGTTTGGCTTCTCTGTGCGCTGGTCAGAGCGGAGACACCTTCGTGCG 120
    |||||
Db 329 GAGGCTCCAAACGTTTGGCTTCTCTGTGCGCTGGTCAGAGCGGAGACACCTTCGTGCG 388
    |||||

Qy 121 CAGGAGTTGGAGCAATTCAGGAGCTCAAGAGACCTCAAGAGACCTTCCTCCGGTGA 180
    |||||
Db 389 CAGGAGTTGGAGCAATTCAGGAGCTCAAGAGACCTTCCTCCGGTGA 445
    |||||

Qy 181 GGGGGGCTCTCGGAGATCTCACCGCTTCTCCAAAGCTTCTCGATGCCACCTGTT 240
    |||||
Db 446 GGGGGGCTCTCGGAGATCTCACCGCTTCTCCAAAGCTTCTCGATGCCACCTGTT 505
    |||||

Qy 241 GGGACGCGTGGGGCGCACGAGCCGCGCTTGGCGCGCTGACGCTGTGGAAACCTATTTC 300
    |||||

```

```

Db 506 GGGACGCGTGGGGCGCACGAGCCGCGCTTGGCGCGCTGACGCTGTGGAAACCTATTTC 565
    |||||
Qy 301 TCGGAGGCTGCTGGCGACTGCAGAGCGCGTGGGACGAGCCCGGAGATCACTGGCTTC-T 359
    |||||
Db 566 TCGGAGGCTGCTGGCGACTGCAGAGCGCGTGGGACGAGCCCGGAGATCACTGGCTTC-T 625
    |||||
Qy 360 TCGCACCGCAACCCCTGGACCTGGAGCCCGCGCTGCCACCCGCGAGCGGGTGATCTCTGC 419
    |||||
Db 626 TCGCACCGCAACCCCTGGACCTGGAGCCCGCGCTGCCACCCGCGAGCGGGTGATCTCTGC 685
    |||||
Qy 420 CCACCCGAGAGGAGCGCTCTTTCTGCGCTGCGGGCGCCCTCTCATCATCACAGTCTGG 479
    |||||
Db 686 CCACCCGAGAGGAGCGCTCTTTCTGCGCTGCGGGCGAG-CTCTCATCATCACAGTCTGG 744
    |||||
Qy 480 AGGCTCAGAGCGCTGCGCTGCTGCTGACGCGCTTCTGTATCCCAAGG 521
    |||||
Db 745 AGGCTCAGAGCGCTGCGCTG-CTGACGCCCTTCTGTATCCCAAGG 785
    |||||

RESULT 3
BX283317
LOCUS 519 bp mRNA linear EST 04-MAR-2003
DEFINITION BX283317 NIH_MGC_7 Homo sapiens cDNA clone IMAGE958E04796 ;
IMAGE:3942507, mRNA sequence.
ACCESSION BX283317
VERSION BX283317.1 GI:28847771
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 519)
Ebert, L., Heil, O., Hennig, S., Neubert, P., Partsch, E., Peters, M.,
Radelof, U., Schneider, D. and Korn, B.
Human Unigeneset - RZPD3
Unpublished (2003)
Contact: Ina Rolfs
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany
RZPD; IMAGE958E04796.
RZPDLIB; I.M.A.G.E. cDNA Clone Collection;
Human Unigeneset - RZPD3 (RZPDLIB No.972)
http://www.rzpd.de/CloneCards/cgi-
bin/showLib.pl.cgi/response?libNo=972 Contact: Ina Rolfs
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Heubnerweg 6, D-14059 Berlin, Germany
Tel: +49 30 32639 101
Fax: +49 30 32639 111
www.rzpd.de
This clone is available royalty-free from RZPD;
contact RZPD (clone@rzpd.de) for further information. Seq primer:
PCMV-M13u, Primer sequence: CGTTGTAAACGAGCGCCAGT.
FEATURES
    source
        1..519
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="IMAGE958E04796 ; IMAGE:3942507"
            /tissue_type="small cell carcinoma"
            /cell_line="MGC3"
            /lab_host="DH10B (phage-resistant)"
            /clone_lib="NIH_MGC_7"
            /note="Organ: lung; Vector: pOTB7; Site 1: XhoI; Site 2:
            EcoRI; cDNA made by oligo-dt priming. Directionally
            cloned into EcoRI/XhoI sites using the following 5',
            adaptor: GGACGAG(G). Size-selected >500bp for average
            insert size 1.8kb. Library constructed by Ling Hong in
            the laboratory of Gerald M. Rubin (University of
            California, Berkeley) using ZAP-cDNA synthesis kit
            (Stratagene) and Superscript II RT (Life Technologies)".
ORIGIN

```

```
Query Match      38.2%; Score 432.2; DB 5; Length 519;
Best Local Similarity 98.9%; Pred. No. 3.2e-79;
Matches 434; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 694 AGGAGACAGCGTCCCCAGTTCTGTGCTTCCCGCGCTACGAGACAGCCGCGCAGATGA 753
Db 17 AGGAGACAGCGTCCNTCCNTCTGTGCTTCCCGCGCTACGAGACAGCCGCGCAGATGA 76
QY 754 GCTGTCCTGTCGCGCGCGCGCGCGTGTGCGCTGTGGAACGTACAGCCGCGCTGGTG 813
Db 77 GCTGTCCTGTCGCGCGCGCGCGCGTGTGCGCTGTGGAACGTACAGCCGCGCTGGTG 136
QY 814 GCTATGACAGTACGCGACCGCGCGCGCTACTCCCGCGGTGCTGCGCGCCGGAAGG 873
Db 137 GCTATGACAGTACGCGACCGCGCGCGCTACTCCCGCGGTGCTGCGCGCCGGAAGG 196
QY 874 GCTGGCGGCTCTCTGAGCGGACCGGGTTCGCTGGAGGAGACGACCCGCGCGGTGAGGC 933
Db 197 GCTGGCGGCTCTCTGAGCGGACCGGGTTCGCTGGAGGAGACGACCCGCGCGGTGAGGC 256
QY 934 CCGGGGCTTCCCTGAACCTCCAGGCGACCGCCCTCCCGCCACGCTGCCACCGGACC 993
Db 257 CCGGGGCTTCCCTGAACCTCCAGGCGACCGCCCTCCCGCCACGCTGCCACCGGACC 316
QY 994 TTCGCGCGGCGCATCCAGAGCGCTGTGTCACCGTTCACAGCAGCGCGCTGGAGCGCG 1053
Db 317 TTCGCGCGGCGCATCCAGAGCGCTGTGTCACCGTTCACAGCAGCGCGCTGGAGCGCG 376
QY 1054 CCCACGCGCCAGGCGCCCTCGAGGGTGTGCTGAGTCTGTGCGCGACCCCGACGACGA 1113
Db 377 CCCACGCGCCAGGCGCCCTCGAGGGTGTGCTGAGTCTGTGCGCGACCCCGACGACGA 436
QY 1114 GCAGTGAGCGCGAGGATCC 1132
Db 437 GCAGTGAGCGCGAGGATCC 455
```

```
RESULT 4
LOCUS      CA311857      599 bp      mRNA      linear      EST 04-NOV-2002
DEFINITION UI-CF-FNO-afe-d-10-0-UI.s1 UI-CF-FNO Homo sapiens cDNA clone
            UI-CF-FNO-afe-d-10-0-UI 3', mRNA sequence.
ACCESSION  CA311857
VERSION     CA311857.1 GI:24529955
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 599)
AUTHORS   Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE     Normalization and subtraction: two approaches to facilitate gene
            discovery
JOURNAL   Genome Res. 6 (9), 791-806 (1996)
MEDLINE   97044477
PUBMED    8889548
COMMENT   Contact: McCray, PB
            McCray Lab
            University of Iowa
            2024 University of Iowa Med Labs, Iowa City, IA 52242, USA
            Tel: 319 356 4866
            Fax: 319 356 7171
            Email: paul-mccray@uiowa.edu
            Tissue Procurement: Dr. M. J. Welsh, University of Iowa
            cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
            cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
            DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
            Clone Distribution: Researchers may obtain clones from Research
            Genetics (www.regen.com) or from Open Biosystems
            (www.openbiosystems.com).
            Seq primer: M13 FORWARD
            POLYA=Yes.
            Location/Qualifiers
```

```
1. 599
/organism="Homo sapiens"
/db_xref="mRNA"
/db_xref="taxon:9606"
/clone="UI-CF-FNO-afe-d-10-0-UI"
/tissue type="Human Lung Epithelial cells"
/lab host="DH10B (Life Technologies) (T1 phage resistant)"
/clone lib="UI-CF-FNO"
/note="Organ: Lung; Vector: pT73-Pac (Pharmacia) with a
modified polylinker; Site1: EcoR I; Site 2: Not I;
UI-CF-FNO is a subtracted cDNA library derived from two
normalized Human lung epithelial cell libraries (EN1 and
DU1) The library was subtracted according to according to
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
1996. For additional information, contact:
bento-soares@uiowa.edu
TAG TISSUE=Human Lung Epithelial Cell Lines untreated LPS
6hr to LPS 24h
TAG LIB=UI-CF-FNO
TAG_SEQ=CTGCTCAGGT"

ORIGIN
Query Match      37.9%; Score 429; DB 6; Length 599;
Best Local Similarity 100.0%; Pred. No. 1.5e-78;
Matches 429; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 704 GGTCCCGAGTTCTGTGCTTCCCGCGCTACGAGACAGCCGCGCAGATGAGTGTCCGTG 763
Db 569 GGTCCCGAGTTCTGTGCTTCCCGCGCTACGAGACAGCCGCGCAGATGAGTGTCCGTG 510
QY 764 CCCGCGGGCGCGCGTGTGGAACGTGAGACCGCGCTGTGGCTATGCAAGG 823
Db 509 CCCGCGGGCGCGCGTGTGGAACGTGAGACCGCGCTGTGGCTATGCAAGG 450
QY 824 TACGCGACCGCGCGGCTTACTCCCGCGTGTGCTGCGCGCGAAGGCTGGCGCT 883
Db 449 TACGCGACCGCGCGGCTTACTCCCGCGTGTGCTGCGCGAAGGCTGGCGCT 390
QY 884 CTCCTGAGCGGACGGGTTCCGTGGAGAGACAGCCGCGGGTGAGCCCGGGGCTTC 943
Db 389 CTCCTGAGCGGACGGGTTCCGTGGAGAGACAGCCGCGGGTGAGCCCGGGGCTTC 330
QY 944 CTTGAACCTCCAGGCGCACCGCCCTCCCGCGCGTGTGCGCGCGCGC 1003
Db 329 CTTGAACCTCCAGGCGCACCGCCCTCCCGCGCGTGTGCGCGCGCGC 270
QY 1004 GCCATCCAGACCGCTGTGCACCGTTCACGCGAGGGCCCTGGAGCGCGCGCCACGCGCG 1063
Db 269 GCCATCCAGACCGCTGTGCACCGTTCACGCGAGGGCCCTGGAGCGCGCGCCACGCGCGC 210
QY 1064 CAGGCGCGCTCGAGGTTGCTGACTCTGTGCGGACCCCGACGACGAGGAGTGAGG 1123
Db 209 CAGGCGCGCTCGAGGTTGCTGACTCTGTGCGGACCCCGACGACGAGGAGTGAGG 150
QY 1124 CGAGGATCC 1132
Db 149 CGAGGATCC 141

RESULT 5
LOCUS      AL042759/c      751 bp      mRNA      linear      EST 04-SEP-2003
DEFINITION DKFPz434C0722.s1 434 (synonym: htes3) Homo sapiens cDNA clone
ACCESSION  AL042759
VERSION     AL042759.2 GI:5935566
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 751)
AUTHORS   Blum,H., Bauersachs,S., Mewes,H.W., Gassenhuber,J. and Wiemann,S.
```

EST (Blum, et al.) Unpublished (1999) On Jul 8, 1999 this sequence version replaced gi:5422206. Contract: MIPS MIPS		Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany This is the 3' sequence of the clone insert Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de; sequenced by LMU (Ludwig Maximilians University, Munich/Germany) within the cDNA sequencing consortium of the German Genome Project. zi sequence also available. This clone (DKFZp34C0722) is available at the RZPD in Berlin. Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@zpd.de.	
FEATURES		Location/Qualifiers	
source		1..751	
		/organism="Homo sapiens"	
		/mol_type="mRNA"	
		/db_xref="taxon:9606"	
		/clone="DKFZp34C0722"	
		/tissue_type="testis"	
		/dev_stage="adult"	
		/lab_host="DH10B"	
		/clone_lib="434 (synonym: htss3)"	
		/notes="Vector: pSport1; Site_1: NotI; Site_2: SalI"	
ORIGIN			
Query Match		37.8%; Score 427.4; DB 1; Length 751;	
Best Local Similarity		99.8%; Pred. No. 3.3e-78;	
Matches		428; Conservative 0; Mismatches 1; Indels 0; Gaps 0;	
QY	704	GGTCCCAGTTCGTGCTTCCCGCGCTACGAGACAGCGCGCAGATGAGTGTCCGTG	763
DB	637	GGTCCCAGTTCGTGCTTCCCGCGCTAAGAGACAGCGCGCAGATGAGTGTCCGTG	578
QY	764	CCGCGGGGCGCGCGTGCCTGTTGGAAACGTCAGACCGCGGCTGGTGGCTATGCAGG	823
DB	577	CCGCGGGGCGCGCGTGCCTGTTGGAAACGTCAGACCGCGGCTGGTGGCTATGCAGG	518
QY	824	TACGGGACCGGGGGGCTACTCCCGCGGTGCTGTCGGCGCGAAGGGCTGGCGCT	883
DB	517	TACGGGACCGGGGCGGCGCTACTCCCGCGGTGCTGTCGGCGCGAAGGGCTGGCGCT	458
QY	884	CTCTCAGGGGACGGGTTCCGTGGAGAGACGCCCGCGGTGAGGCCCGGGGCTTC	943
DB	457	CTCTCAGGGGACGGGTTCCGTGGAGAGAGACGCCCGCGGTGAGGCCCGGGGCTTC	398
QY	944	CCTGAACCTTCCAGGACACCGCCCTTCCCGCCACCGTGCACCCGACCTTCGCGGGGC	1003
DB	397	CCTGAACCTTCCAGGACACCGCCCTTCCCGCCACCGTGCACCCGACCTTCGCGGGGC	338
QY	1004	GCCATCCAGACCGCTGTGACCGTACACGACGAGGCGCTTGAGCGGCGCCACGCGCGC	1063
DB	337	GCCATCCAGACCGCTGTGACCGTACACGACGAGGCGCTTGAGCGGCGCCACGCGCGC	278
QY	1064	CAGGGCGGCGCTCGAGGTCGTTGACTGTCGCCGCCACCCACGACGAGCAGTGCAGCG	1123
DB	277	CAGGGCGGCGCTCGAGGTCGTTGACTGTCGCCGCCACCCACGACGAGCAGTGCAGCG	218
QY	1124	CGAGGATCC	1132
DB	217	CGAGGATCC	209
RESULT 6		623 bp mRNA linear EST 20-SEP-2000	
BE793330		BE793330	
LOCUS		LOCUS	
DEFINITION		DEFINITION	
ACCESSION		ACCESSION	
VERSION		VERSION	

EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 623)
REFERENCE NIH-MGC <http://mgs.nci.nih.gov/>.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: DCTD/DTP
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LLCM796 row: e column: 04
High quality sequence start: 28
High quality sequence stop: 616.
Location/Qualifiers
1..623
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3942507"
/issue_type="small cell carcinoma"
/cell_line="MGC3"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_7"
/note="Organ: lung; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5',
adaptor: GGCACGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."

ORIGIN

Query Match	37.2%;	Score	421.2;	DB 2;	Length	623;		
Best Local Similarity	99.1%;	Pred.	No. 6.2e-77;					
Matches	434;	Conservative	0;	Mismatches	3;	Indels	1; Gaps	1;
QY	695	GGGAGCAGCGGTCCCCAGTTCTGTGCTTCCTCCGCGCTACGAGCAGCGCGCAGATGAG	754					
Db	41	GGAGCCAGCGGTCCCCAGTTCTGTGCTTCCTCCGCGCTACGAGCAGCGCGCAGATGAG	100					
QY	755	CTGTCCGTCCCGCGGGGCGCGTGCCTGTTGAAACCGTACACCGCGCTCGTGG	814					
Db	101	CTGTCCGTCCCGCGGGGCGCGTGCCTGTTGAAACCGTACACCGCGCTCGTGG	160					
QY	815	CTATGCAAGTACGGACCGGGGGCGCTACTCCCGCGGTGCTGTCGGCCGGAAGGG	874					
Db	161	CTATGCAAGTACGGACCGGGGGCGCTACTCCCGCGGTGCTGTCGGCCGGAAGGG	220					
QY	875	CTGGGGCGCTCTCTGAGCGGACGGGTTCCGTGGAGGAGACGACCCGGCGGTGAGGCC	934					
Db	221	CTGGGGCGCTCTCTGAGCGGACGGGTTCCGTGGAGGAGACGACCCGGCGGTGAGGCC	280					
QY	935	CGGGGCTTCCTGAAACCTCCAGGCCACCGCCCTCCGCCCAACCGTGCACCCGACCT	994					
Db	281	CGGGGCTTCCTGAAACCTCCAGGCCACCGCCCTCCGCCCAACCGTGCACCCGACCT	340					
QY	995	TGCGCGGGCGCATTCAGAGCGCTGCTGCACCGTACACGAGGGCCCTTGAGCGGGCGC	1054					
Db	341	TGCGCGGGCGCATTCAGAGCGCTGCTGCACCGTACACGAGGGCCCTTGAGCGGGCGC	400					
QY	1055	CCACGGCGCAGGGCGCCCTCGAGGGTGCCTGCTGCGCGCACCCACGACGGAG	1114					
Db	401	CCACGGCGCCA-GGGCGCCCTCGAGGGTGCCTGCTGCGCGCACCCACGACGGAG	459					
QY	1115	CAGTGAGCGCGAGGATCC	1132					

[illegible]

```

Db      460 CAGTGAGCGGAGGATCC 477
|||||
BG328824      473 bp mRNA linear EST 27-FEB-2001
602427426F1 NIH_MGC_15 Homo sapiens cDNA clone IMAGE:4546878 5',
mRNA sequence.
BG328824
BG328824.1 GI:13135262
EST.
Homo sapiens (human)
Homo sapiens
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 473)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: NIH Intramural Sequencing Center
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCW1234 row: c column: 07
High quality sequence stop: 471.
FEATURES
Location/Qualifiers
1..473
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4546878"
/tissue_type="adenocarcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_15"
NOTE="Organ: colon; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)"
ORIGIN
Query Match 37.2%; Score 421; DB 4; Length 473;
Best Local Similarity 97.4%; Pred. No. 6.7e-77;
Matches 449; Conservative 0; Mismatches 10; Indels 2; Gaps 2;
QY 672 GCGGGGAGGAGCGCGCTCCCTAGGAGCAGCGGTCCCAGTTCTGTGCTTCCCGCGCT 731
Db 1 GCGGGGAGGAGCGCGCTCCCTAGGAGCAGCGGTCCCAGTTCTGTGCTTCCCGCGCT 60
QY 732 ACGAGAGCAGCGCGCAGATGAGTGTTCGTCGCCCGCGGCGCGCGTGCAGTGTGG 791
Db 61 ACGAGAGCAGCGCGCAGATGAGTGTTCGTCGCCCGCGGCGCGCGTGCAGTGTGG 120
QY 792 AAACGTACAGCGCGCTGTGGCTATGAGTACGGACCGGCGGCGCTACTTCCCG 851
Db 121 AAACGTACAGCGCGCTGTGGCTATGAGTACGGACCGGCGGCGCTACTTCCCG 180
QY 852 CGGTCTCTGCGCGCGGAAGGCTGGCGCTCTCTGAGCGGAGCGGGTTCGTGGAG 911
Db 181 CGTCTCTGCGCGCGGAAGGCTGGCGCTCTCTGAGCGGAGCGGGTTCGTGGAG 240
QY 912 GAGACGACCGCGGCTGAGGCGCGGGCTTCCCTGAACCCCTCCAGGCCACCGCCCTC 971
Db 241 GAGACGACCGCGGCTGAGGCGCGGGCTTCCCTGAACCCCTCCAGG-CACCGACCTC 299

```

```

QY 972 CCCCCACCGTCCCGACCGACCTTCGCGGGGCGCCATCCAGAGCCGCTGTCACCGTCA 1031
Db 300 CCCCCACCGTCCCGACCGACCTTCGCGGGGCGCCATCCAGAGCCGCTGTCACCGTCA 359
QY 1032 CACGAGCGGCGCTGAGCGGGGCGCCACGCGCGCGCGCGCGCTCGAGGGTGCCTGGACT 1091
Db 360 CACGAGCGGCGCTGAGCGGGGCGCCACGCGCGCGCGCGCGCTCGAGGGTGCCTGGACT 419
QY 1092 CTGTGCGGACCCACGACGAGCGAGTGTGAGCGCGAGGATCC 1132
Db 420 CTGTGCGG-ACCCACGACGAGCGAGTGTGAGCGCGAGGATCC 459

RESULT 8
AK009605      1571 bp mRNA linear HTC 03-APR-2004
Mus musculus adult male tongue cDNA, RIKEN full-length enriched
library, clone:2310034C04 product:2310034C04RIK PROTEIN (SNX28)
homolog [Mus musculus], full insert sequence.
ACCESSION AK009605
VERSION AK009605.1 GI:12844500
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1
Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
9279253
10349636
2
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
11042159
3
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yanamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multipillar sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
11076861
4
The RIKEN Genome Exploration Research Group Phase II Team and the
PANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)
5
The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 1571)
Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Arai, A., Aono, H.,
Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Fukunishi, Y.,
Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hiramoto, K.,
Hiraoka, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Izawa, M.,
Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K.,
Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C.,

```



```
/clone lib="NCI CGAP_SG2"
/Note="Organ: salivary gland; Vector: pCMV-SPORT6; Site 1:
NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo
dT. Average insert size 1.3 kb. Constructed by Life
Technologies. Note: this is a NCI CGAP Library."
ORIGIN
Query Match 33.2%; Score 376.2; DB 4; Length 904;
Best Local Similarity 70.9%; Pred. No. 1.5e-67;
Matches 543; Conservative 0; Mismatches 218; Indels 6; Gaps 3;
QY 13 CCCCCGATACCCAGTTTCAGTCAAGGGGCGAGCCCTGGTGCAGATCAAGAGGCTCCAAAC 72
DB 1 CCAAGACACCCAGATATAGCCCATGCTAGACCTTGGTGCAATGGACCGACTCCAGAC 60
QY 73 GTTTCCTTCTCTGCGCTGTCAGACGGCGAGCGACACTTTCGTGCGCAGAGTTGGGA 132
DB 61 ATTTCCTTCTCCGTGTGCTGTCAGACAAACAGTGACACATTTGTACGGAGGAGCTGGGA 120
QY 133 CCAATTTCAGGAGCTCAAGAAACCTTCAAGAGACCTTCCCGTGGAGCGGGCTGCT 192
DB 121 TCAGTTCAGGCGAGCTCCAGAAACCTTAAAGAAACCTTCCAGTGGAGCGGCGCTGCT 180
QY 193 GCGGAGATCTGACCGCTTCTCCAAAGCTTCTCGATGCACCACTGTGGGACGCGTGG 252
DB 181 ACGGAGATCTGAACAGTTCTTCCAGCTTCTGATGCTCCATTTGCTGACACGTCGGGG 240
QY 253 GCGCACGAGCCGCGGCTGCGCGCTGCGAGCTGTGTGGAACCTATTTCGGAGGCTGCT 312
DB 241 GCATATCTGCTCAGGACTGGTACGTTTGGGCTGCTGGACACATATGTACAGGCATTGCT 300
QY 313 GCGGACTCAGAGCGGTGGACGAGGCCGACGATCACTGCTTCTTCGACCGCAACC 372
DB 301 GGCACACCTCAGAACACATATTGAGGAGTTCAGCACTTTCACCGCTTCTTGTACCCAAACC 360
QY 373 CTGTGACCTGAGCCGCGCTCCACCGGACGCGGCTGATCTGCGCCACCCACAGAGA 432
DB 361 TGTGATCTGGAGCCCATGCTGCTCTGGAGCTGTGATCTTCCCTACACAGAGG- 419
QY 433 GCAGCTCTTTCTCGGCTGCGGGCGGCTCTCCATCCACAGTCTGGAGGCTCAGAGCT 492
DB 420 --AGCCCTTATCCCAACCCAGAGCGAGCCTTGACATTCACAGCTGGAGGCTCAGAGCAT 477
QY 493 GCGCTGCTGACGCTTCTGTACCCAGACAGCGGGATAGGCTTTTCAGGCGCAGGC 552
DB 478 TCCCTGTGTACAGTCTTTCACACTTTCACATTAAGAGACAGACACCTTTCACACCAAGGC 537
QY 553 CCAGGAGAGCTGGAGCTGCTGCTGCGCACCCCTCAGGCTGGTGGTGGAGAACGA 612
DB 538 TCAAGAAATCTGGACATATTAATAGCATCTTCAAGCTGTGTGGTGGTGGAGAACAA 597
QY 613 AGACGGCAGACCGCTGTTTCCAGCGCTTACCTGGAGGAGCGGCCCG-GGCCCAAG 671
DB 598 GGATCAGCAGGTAGCTGTTTCCAGCTTCTTACCTGGAGGAGGTAGCAACGTCGCCAAG 657
QY 672 GCCGGAGGAGGCGGCTTCCCTAGGAGCAGCGGTCCCGAGTTCTGTGCT--TCCCGGCG 729
DB 658 GCCAGAGTTCAGGCGCTGCTGTTGCAAGGAAGTGGGAGGCGATTCTGGAATACCCCAAGC 717
QY 730 CTACGAGAGCGCCGCGAGATGAGCTGTCGTCGCCCGGGGGGCG 776
DB 718 CTACGAGGCGACGCGCTCTGATGAGCTAATCCGGGGGCTTCAGGGGC 764
```

```
RESULT 10
BI409553
LOCUS 602963963F1 NCI_CGAP_Lu33 Mus musculus cDNA clone IMAGE:5119670 5',
DEFINITION mRNA sequence.
ACCESSION BI409553
VERSION BI409553.1 GI:15170476
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
```

```
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: c9apbs-remail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11291 row: m column: 15
High quality sequence start: 7
High quality sequence stop: 724.
Location/Qualifiers
1..726
/organism="Mus musculus"
/mol_type="mRNA"
/strain="CZECH II"
/db_xref="taxon:10090"
/clone="IMAGE:5119670"
/tissue_type="pooled lung tumors"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NCI CGAP Lu33"
/Note="Organ: lung; Vector: pT7T3D-Pac (Pharmacia) with a
modified polylinker; Site 1: NotI; Site 2: EcoRI; 1st
strand cDNA was prepared from mRNA obtained from pooled
lung tumors with a Not I - oligo(dT) primer [5'
TGTTCACCAATCTGAATGGAGCGCGGCTCTGTTTTTTTTTTTTTTT 3'].
Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not
I and Eco RI sites of the modified pT7T3 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo."
```

ORIGIN

```
Query Match 30.7%; Score 347; DB 4; Length 726;
Best Local Similarity 70.9%; Pred. No. 1.5e-61;
Matches 518; Conservative 0; Mismatches 205; Indels 8; Gaps 4;
QY 12 GCCCCCCGATACCCAGTTTTCAGTCAAGGGGCGAGCCCTGGTGCAGATCAAGAGGCTCCAAA 71
DB 3 GGGCCCAAGACACCCAGTATGCCCATGTTGTAGCCTTGTGCA-ATGGACCGACTCCAGA 61
QY 72 GCTTTGCTTCTCTGTGCGCTGGTTCAGACGGGAGC-GACACCTTCTGTGGCAGGAGTTGG 130
DB 62 CATTTGCCCTTCTCGGTGTGCTGGTTCAGACAAACAGGTGACACATTTGTACGGAGGAGCTGG 121
QY 131 GACCAATTCAGCAGCTCAAGAGACCTTCAGGAGACCTTCCCGTGGAGCGGCGCTG 190
DB 122 GATGAGTTTCAGGAGCTC---CAGACCTTAAAGAAAACCTTCCCAGTGGAGCGAGCCTG 178
QY 191 CTGGGAGATCTGACCGGCTTCTCCAAAGCTTCTCGATGCACCACTGTGGGAGCGCGTG 250
DB 179 CTACGGAGATCTGAACAAGTTCTTCCCAAGCTTCTGATGCTCCATTGCTGACACGTCGG 238
QY 251 GGGCGCAGACGCGCGCTGCGCGCTGAGCTGTTGGAAAACCTATTCTCGGAGGCTG 310
DB 239 GGGCATACTGTCGAGGACTGCTACGTTTGGGGTGTCTGGACACCTATGTACAGGCATTG 298
QY 311 CTGGCATCTGAGAGCGGTGGCAGGAGCCCGACGATCACTGGCTTCTTCGACCGGAA 370
DB 299 CTGGCAACTCAGAACACATATTGAGGAGTTTACGACTTCACGGCTTCTTTGTACCCAAA 358
QY 371 CCCCTGGACCTGGAGCCCGCGCTGCCACCGGCGAGCGGGTGATCCTCCCAACCCAGAG 430
DB 359 CCTCTGGATCTGGAGCCCATCTGCCTCTGCGAGCCTGGTGATCCTGCCTACACAGAG 418
```

Qy	431	GAGCAGCCTCTTTCTCGCGCTCGCGCGCGCTCTCCATCCACAGTCTGGAGGCTCAGAGC	490
Db	419	G --- AGCCCTTATCCCAACCCAGAGGCGGCTTGACATTCACAGCTGGAGGCTCAGAGC	475
Qy	491	CTGCGCTGCCTGCAGCCCTTCTGTATCCAGAGACACGGGATAGGCTTTTTCAGGCGCAG	550
Db	476	ATTCCCTGTGTACAGCCTTCCACACTCTTGACATATAGAGACAGACTTTCACACCCAG	535
Qy	551	GCCCAGGAGAGCCTGGAGCTGCTGCTCGCGGACCCCTCAGGCTGTGTGGCTGTGGAGAAC	610
Db	536	GCTCAAGAAATTCGGACATATTAACGACATCTTCAGGCTGTGTGGCTGTGGAGAAC	595
Qy	611	GAAGACCGGACAGCGCTGTGTTTCAGCGCCCTACTCGAGAGAGCGGCCCCCGGGCCAA	670
Db	596	AHGGATCAGCAGGTAGCCCTGTGTTTTCAGCTCCCTACTCGAGAGAGGTAGCAACGTGCCAA	655
Qy	671	GGCCGGGAGGAGCGCCGTCTCTAGGAGACAGCGGTCCCAGTTCTGTGCTTCCCAGCGCC	730
Db	656	GGCCAGGAGTCAGGCTGGCTTGTCAAGGAAGTGGGAGGCGAGTTCTGCACTACCCAGGCC	715
Qy	731	TACGAGAGCAG	741
Db	716	TACGAGGCGAG	726

RESULT	11
BE751524	
LOCUS	544 bp mRNA linear EST 25-APR-2001
DEFINITION	203692 MRC 2BOV Bos taurus cDNA 5', mRNA sequence.
ACCESSION	BE751524
VERSION	BE751524.1 GI:10165516
KEYWORDS	EST.
SOURCE	Bos taurus (cow)
ORGANISM	Bos taurus
Eukaryota;	
Mammalia;	Chordata; Craniata; Vertebrata; Euteleostomi;
Bovinae;	Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
1 (bases 1 to 544)	
Smith,T.P.L., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,R.T.,	
Cassas,E., Wray,J.E., White,J., Cho,J., Fahrrenkrug,S.C.,	
Bennett,G.L., Heaton,M.P., Laegreid,W.W., Rohrer,G.A.,	
Chitko-McKown,C.G., Ferrea,G., Holt,I., Karamycheva,S., Liang,F.,	
Quackenbush,J. and Keele,J.W.	
Title	Sequence evaluation of four pooled-tissue normalized bovine cDNA libraries and construction of a gene index for cattle Genome Res. 11 (4), 626-630 (2001)
JOURNAL	

MEDLINE PUBMED COMMENT	TEXT
21180013	USDA, ARS, US Meat Animal Research Center PO Box 166, Clay Center, NE 68933-0166, USA Tel: 402 762 4366 Fax: 402 762 4390 Email: smith@email.marc.usda.gov Single pass sequencing. Bases called and alt trimmed with phred v0.980904.s. Vector identified by cross_match with the -minscore 18 and -minmatch 12 options.
11282978	Contact: Smith TPL

```

PCR Primers
FORWARD: AGGAAACAGCTATGACCAT
BACKWARD: GTTTTCCAGTCAGCAGC
Plate: 43 row: L column: 21
Seq primer: ATTAGTGACACTATAG.
Location/Qualifiers
1..544
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/tissue_type="pooled"
/lab_host="DH10B"
/clone_lib="MARC 2B0V"
/note="Vector: pCMV SPORT6; Site 1: NotI; Site 2: SalI;
Library made from pooled tissue from testis, thymus,

```

ORIGIN	Query Match	29.6%;	Score 334.8;	DB 2;	Length 544;	
	Best Local Similarity	77.7%;	Pred. No. 4.9e-59;			
	Matches 418;	Conservative	0;	Mismatches 117;	Indels 3;	Gap# 1;
Qy	226	CGATGCACCACTGTTGGAGCGCTGGGGGCGACAGAGCGCGCGCTTGCGCGCTTGACGCT	285			
Db	10	CGATGCACCTGGCTGCTGACCAAGACCGCAGCGGCGAGAGCGCTGCGCGCTTGCGAGCT	69			
Qy	286	GTTTGGAAACCTATTCTCGAGGCTGCTGGCGACTGAGAGCGCGTGCAGAGCGCGTGCACGAGGAGCCGAC	345			
Db	70	GCTGGAGGACTACTGAGGGGACTGCTGGCGCGCGGGGAGCGTGTGTCTCCGCGACCTTGGGA	129			
Qy	346	GATCACTGGGCTTCTTCGACACCGCAACCCCTGGAGCTGGAGCGCGCGCTGCGACACCCCGGCGAG	405			
Db	130	GCTCACTGGCTTCTTACACCAACCTCTGGACCTGGAGCGCTATGCTGCCACTGGCAG	189			
Qy	406	CCGGGTGATCTGCTGCCACCCCGAGAGGAGCGCTCTTTTCTCGCGCTGCGGGCGCGCTTCT	465			
Db	190	CTTGCTGATCTGCCACACCCCGCAGAGCGCTCGC--CAGGCGCCCTCAGGCGAGCCCTGC	246			
Qy	466	CATCCACAGTCTGGAGGCTCAGAGCGCTGCGCTGCTGCGAGCGCTTCTGTATCCAGGACAC	525			
Db	247	GGTCCACACCTCGAGGCGCCAGAGCGCTGCACCTGCGCTGCAGACCTTCAGCACCCAGGACAC	306			
Qy	526	GCGGGATAGGCTTTTTCAGCGCAGCGACCCAGGAGAGCCTGGAGCTGCTGCTGCGGGCACCC	585			
Db	307	GCGGGCGAGCCCTTCCAGCTCAGCGCCAGGAGCGCTTGGATGCTGCTGACAGACCC	366			
Qy	586	CTCAGGCTGGTGGCTGGTGGAGAACCGGACAGCGCGCTGCTGCTTCCAGCGCGCTTA	645			
Db	367	CTCAGGCTGGTGGCTGGTGGAGAACCGGACAGCGCGCTGCTGCTTCCAGCTCCCTA	426			
Qy	646	CCTGGAGGAGGCGGCGCGCGGCCAACGCGCGGAGGAGCGCCCTGCTTCCAGGAGGAGCGG	705			
Db	427	CCTGGAGGAGGCGGCGCGGCCAACCCAGGAGGAGGATGGGGGACAGCTCTCGGGGATAAGCGG	486			
Qy	706	TCGCCAGTTCGTGCTTCCGCGGCTTACAGAGCAGCGCGGAGATGAGCTGCTGCTG	763			
Db	487	GCCCAAGTTCGTGCTTCCGCTTATGAGAGCAGCGCGGCTGATGAGCTGCTGCTG	544			
RESULT 12	602359714f1 NCI_CGAP_Mam1 Mus musculus cDNA clone IMAGE:4488315 5',	773 bp	mus	linear	EST 13-FEB-2001	
LOCUS	mus musculus (house mouse)					
DEFINITION	mus musculus (house mouse)					
ACCESSION	CG247058					
VERSION	CG247058.1					
KEYWORDS	EST					
SOURCE	mus musculus (house mouse)					
ORGANISM	mus musculus					
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
AUTHORS	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.					
TITLE	1 (bases 1 to 773)					
JOURNAL	NIH-MGC http://mgc.nci.nih.gov/					
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC)					
	Unpublished (1999)					
	Contact: Robert Strausberg, Ph.D.					
	Email: cgapbs-r@mail.nih.gov					
	Tissue procurement: Gilbert Smith, Ph.D.					
	cDNA Library Preparation: Life Technologies, Inc.					
	DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)					
	Clone distribution: MGC clone distribution information can be					
	found through the I.M.A.G.E. Consortium/LLNL at:					
	http://image.llnl.gov					
	Plate: LLAM10334 row: k column: 04					
	High quality sequence stop: 663.					
	Location/Qualifiers					
FEATURES						

Db 499 GGCTCAGAGATTCCCTGTGTACAGCCTTTCCACACTTTTGACATAGAGACAGACCTTT 558
 QY 541 TCAGGGCCAGCCAGAGAGCTTGGACGTGCTGCTGCGGCAC - CCTCAGGCTGG-TG 597
 Db 559 CCACACCAAGGCTCAGAAATTTCTGGCACATTTACTTTACGCACATCTTTCAGGCTGGTTG 618
 QY 598 GCTGTGTGAGAACGAAGA - CCGGCAGACCGCCTGGTTT - CCAGGCGCTTACCT - GGAGGA 654
 Db 619 GCTGTGTGAGAACGAAGATCACGAGGTAGCTGTGTTTCCAGATCCCTACCTGGGAGGA 678
 QY 655 GCGCGCCCGGGCCAGGCGGAGGA ----- GCGCGTCTCCTAGGAGCAGCGGTCC 708
 Db 679 GGTAGCAAAAGCTCCAGGCGAGAGTCAAGCGCTGGCTTTTTCACAGGGAATCGGCGAG 738
 QY 709 CCAGTCTCTGTCTTCCCGCCCTACGAGAGCAGCGCGCAGATGAGCTGTCCGTGCGCCGC 768
 Db 739 GCAGTCTGCACCTACCCAGCATACGACGAGTCGCTCTGTAGGAGTATCCGCGCCCGAG 798
 QY 769 GG 770
 Db 799 GG 800

RESULT 14
 CB113007
 LOCUS
 DEFINITION K-EST0154911 L6ChoCK0 Homo sapiens cDNA clone L6ChoCK0-8-B11 5', mRNA linear EST 28-JAN-2003
 mRNA sequence.

ACCESSION CB113007.1 GI:27938814
 VERSION EST.
 KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens
 ORGANISM Homo sapiens

REFERENCE
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 602)
 Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R., Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and Kim,Y.S.

TITLE 21C Frontier Korean EST Project 2001
 JOURNAL Unpublished (2002)
 COMMENT Contact: Kim YS
 Genome Research Center
 Korea Research Institute of Bioscience & Biotechnology
 52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
 Tel: +82-42-860-4470
 Fax: +82-42-860-4409
 Email: yongsung@mail.kribb.re.kr
 Plate: 8 row: B column: 11
 High quality sequence stop: 602.
 Location/Qualifiers
 1..602
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="L6ChoCK0-8-B11"
 /sex="M"
 /cell_line="Cho-CK"
 /lab_host="Top10F"
 /clone_lib="L6ChoCK0"
 /notes="Organ: Liver; Vector: pcNS-D2; Site 1: EcoRI; Site 2: NotI; The poly (A)+ RNA was dephosphorylated with bacterial alkaline phosphatase (BAP) and then decapped with tobacco acid pyrophosphatase (TAP). The decapped intact mRNA was ligated with DNA-RNA linker including EcoRI site by treatment of T4 RNA ligase and the first strand cDNA was synthesized from oligo dt-selected mRNA by priming with dt-tailed vector. The dt-tailed vector was adjusted to have about 60nt. The cDNA vector was circularized with E. coli DNA ligase after digestion of EcoRI which site is also included in vector. An RNA strand converted to a DNA strand by Okayama-Berg method. The obtained cDNA vectors were used for transformation of

ORIGIN

Query Match 28.0%; Score 317.4; DB 6; Length 602;
 Best Local Similarity 99.7%; Pred. No. 2e-55; Indels 0; Gaps 0;
 Matches 318; Conservative 0; Mismatches 1;

QY 1 AGCATGGCAGGCCCCGATACCCAGTTCAGTGCAGGGGAGCCCTGGTGAGATCAA 60
 |||||
 Db 284 AGCATGGCAGGCCCCGATACCCAGTTCAGTGCAGGGGAGCCCTGGTGAGATCAA 343
 |||||
 QY 61 GAGGCTCCAAACGTTTGCCTTCTCTGCGCTGGTTCAGACGGCAGCAGACCTTCGTGCG 120
 |||||
 Db 344 GAGGCTCCAAACGTTTGCCTTCTCTGCGCTGGTTCAGACGGCAGCAGACCTTCGTGCG 403
 |||||
 QY 121 CAGGAGTTGGGACGAATTCAGGAGCTCAAGAGACCTCAAGGAGACCTTCCCGGTGA 180
 |||||
 Db 404 CAGGAGTTGGGACGAATTCAGGAGCTCAAGAGACCTCAAGGAGACCTTCCCGGTGA 463
 |||||
 QY 181 GCGGCGCTGCTGCGGAGATCTGACCGCTTCTCCAAAGCTTCTCGATGCACCACTGTT 240
 |||||
 Db 464 GCGGCGCTGCTGCGGAGATCTGACCGCTTCTCCAAAGCTTCTCGATGCACCACTGTT 523
 |||||
 QY 241 GGGACGCTGGGGCGCACGAGCGCGGCTTGGCGGCTGCGAGCTGTGGAAACCTATTTC 300
 |||||
 Db 524 GGGACGCTGGGGCGCACGAGCGCGGCTTGGCGGCTGCGAGCTGTGGAAACCTATTTC 583
 |||||
 QY 301 TCGGAGGCTGTGGCGCACT 319
 |||||
 Db 584 TCGGAGGCTGTGGCGCACT 602
 |||||

RESULT 15

BQ935073
 LOCUS
 DEFINITION AGENCOURT 8799923 NCI CGAP Co24 Mus musculus cDNA clone IMAGE:6399072 5', mRNA sequence.

ACCESSION BQ935073
 VERSION BQ935073.1 GI:22350456
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus

REFERENCE
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 967)
 NIH-MGC http://mgc.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-r@mail.nih.gov
 Tissue Procurement: The Cepko Laboratory
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLAM13899 row: j column: 01
 High quality sequence stop: 572.
 Location/Qualifiers
 1..967
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="FVB/N"
 /db_xref="taxon:10090"
 /clone="IMAGE:6399072"
 /lab_host="DH10B (T1 phage-resistant)"
 /clone_lib="NCI_CGAP_Co24"
 /note="Organ: colon; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.6 kb. Constructed by Life

FEATURES

source

ORIGIN

Search completed: May 30, 2005, 23:16:24
Job time : 5927 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 28, 2005, 07:23:17 ; Search time 162 Seconds
(without alignments)
885.729 Million cell updates/sec

Title: US-10-621-113-4
Perfect score: 1950
Sequence: 1 MAGPRYPVSQGAALVQIKR.....RRQRPGRGCVDSVPHPTTEQ 371

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A Geneseq_16Dec04:.*
1: Geneseqp1980s:.*
2: Geneseqp1990s:.*
3: Geneseqp2000s:.*
4: Geneseqp2001s:.*
5: Geneseqp2002s:.*
6: Geneseqp2003as:.*
7: Geneseqp2003bs:.*
8: Geneseqp2004s:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1950	100.0	371	8 ADJ57071	Adj57071 Nox enzym
2	1937.5	99.4	376	8 ADJ57073	Adj57073 Nox enzym
3	1934.5	99.2	370	8 ADJ57069	Adj57069 Nox enzym
4	1922	98.6	375	8 ADJ57075	Adj57075 Nox enzym
5	1418	72.7	419	4 ABG24891	Abg24891 Novel hum
6	1418	72.7	726	4 ABG25247	Abg25247 Novel hum
7	309.5	15.9	390	8 ADG36779	Adg36779 Human P47
8	307.5	15.8	390	8 ADG36778	Adg36778 Human P47
9	307.5	15.8	390	8 ADG62962	Adg62962 p47(phox)
10	307.5	15.8	390	8 ADP23795	Adp23795 PRO polyp
11	307	15.7	389	8 ADG36776	Adg36776 Human P47
12	306.5	15.7	390	8 ADRI14450	Adri14450 Human NF-
13	306.5	15.7	976	4 ABG20715	Abg20715 Novel hum
14	306	15.7	389	8 ADG36774	Adg36774 Rat P47PH
15	304.5	15.6	366	8 ABM80885	Abm80885 Tumour-as
16	304.5	15.6	390	2 AAR83825	Aar83825 p47(phox)
17	304.5	15.6	1379	4 ABG20711	Abg20711 Novel hum
18	296	15.2	1619	4 ABG20707	Abg20707 Novel hum
19	278.5	14.3	386	6 AAO16002	Aao16002 Human NAD
20	246.5	12.6	1124	6 AAE38441	Aae38441 Mouse FIS
21	211	10.8	1054	6 ADA55008	Ada55008 Human pro
22	178	9.1	268	8 ABM80892	Abm80892 Tumour-as
23	163	8.4	268	4 AAU17156	Aau17156 Novel sig
24	163	8.4	268	7 ADB93864	Adb93864 Human nov
25	163	8.4	760	6 AAE34861	Aae34861 Human kin

26	161	8.3	940	6 AAE38440	Aae38440 Human FIS
27	158	8.1	204	4 AAU87356	Aau87356 Novel cen
28	158	8.1	204	8 ADI54671	Adi54671 Novel hum
29	153.5	7.9	215	4 AAU68543	Aau68543 Human nov
30	146	7.5	968	4 AAU14174	Aau14174 Human nov
31	143	7.3	134	6 ABUI0058	Abui0058 Human NAD
32	143	7.3	134	6 ABUI0044	Abui0044 Human NAD
33	142.5	7.3	128	4 AAG64031	Aag64031 Human PX
34	142.5	7.3	128	4 AAG64033	Aag64033 Human PX
35	142	7.3	1061	2 AAU87504	Aau87504 Human N-m
36	142	7.3	1212	2 AAU87503	Aau87503 Human N-m
37	141.5	7.3	769	8 ADN21559	Adn21559 Bacterial
38	139.5	7.2	559	1 AAP70196	Aap70196 Sequence
39	139.5	7.2	560	3 AAY92020	Aay92020 Human mul
40	139.5	7.2	560	6 ABB82988	Abb82988 Human mul
41	139.5	7.2	560	8 ADH11597	Adh11597 Human bon
42	139.5	7.2	1081	3 AAB26240	Aab26240 Human N-m
43	139.5	7.2	1081	5 ABB56523	Abb56523 Human NMD
44	139.5	7.2	1081	6 ABU61440	Abu61440 Human N-m
45	139.5	7.2	1081	7 AAE39251	Aae39251 Human NMD

ALIGNMENTS

RESULT 1
ADJ57071
ID ADJ57071 standard; protein; 371 AA.
XX
AC ADJ57071;
XX
DT 06-MAY-2004 (first entry)
DE
DE Nox enzyme regulatory protein p41Nox.
XX
KW Nox regulatory protein; p41Nox; NADH oxidase; cytosolic; human.
XX
OS Homo sapiens.
XX
PN WC2004007689-A2.
XX
PD 22-JAN-2004.
XX
PF 16-JUL-2003; 2003WO-US022246.
XX
PR 16-JUL-2002; 2002US-0396170P.
PR 23-AUG-2002; 2002US-0405647P.
XX
(UYEM-) UNIV EMORY.
PA Lambeth JD, Cheng G;
XX
PI WPI; 2004-122927/12.
XX
DR N-PSDB; ADJ57070.
XX
PT New regulatory protein for hydrogenated nicotinamide adenine dinucleotide phosphate oxidase, useful for preparing a medicament for affecting superoxide formation or regulating cellular proliferation in an animal or a human.
XX
PS Claim 1; SEQ ID NO 4; 70pp; English.
XX
CC The present sequence is the protein sequence of a human regulatory protein, denoted p41Nox, for Nox (NADH oxidase) enzymes. The protein and nucleotide sequences of four p41Nox variants have been identified CC
CC ADJ57068-ADJ57075. p41Nox proteins have a mol.wt. of 41 kDa. They include CC
CC a PX domain and 2 SH3 domains and show sequence homology (25% identity) CC
CC with p47phox, a regulatory protein for gp91phox. They function in the CC
CC regulation of cell growth and are therefore implicated in diseases CC
CC involving abnormal cell growth, such as cancer. They may also function in CC
CC innate immune mechanisms of epithelial tissue or other barrier cells, and CC
CC hence may be involved in diseases of diminished ability to fight CC
CC infections or inflammatory conditions. p41Nox regulatory proteins and

CC nucleic acids are useful for preparing a medicament for affecting
 CC superoxide formation or regulating cellular proliferation in an animal or
 CC a human. The regulatory proteins are also useful in drug development,
 CC e.g. screening for drugs that regulate the biological activity of the
 CC proteins, or in assays that relate to assessment of abnormal growth or
 CC cellular proliferation including cancer.

XX Sequence 371 AA;

Query Match 100.0%; Score 1950; DB 8; Length 371;
 Best Local Similarity 100.0%; Pred. NO. 4.1e-169;
 Matches 371; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGPRYPVSVOGAALVQIKRLQTFAPSVRWSGSDTFVRRSWDEFRQLKTKLKEFPVEA 60

DB 1 MAGPRYPVSVOGAALVQIKRLQTFAPSVRWSGSDTFVRRSWDEFRQLKTKLKEFPVEA 60

QY 61 GLLRRSDRVLPKLLDAPLLGRVGRSGRLARLQLLETYSRRLATAERVARSPITIGFFA 120

DB 61 GLLRRSDRVLPKLLDAPLLGRVGRSGRLARLQLLETYSRRLATAERVARSPITIGFFA 120

QY 121 POPLDLEPALPGSRVILPTPEEQPLSRAAGRLSIHSLEAQSLRCLQPFCTQDTRDPFQ 180

DB 121 POPLDLEPALPGSRVILPTPEEQPLSRAAGRLSIHSLEAQSLRCLQPFCTQDTRDPFQ 180

QY 181 AQAQESLDVLLRHPGSMWLVENEDQTAFPPAPYLEEAAPGQREGPSLGSSGPFQFCAS 240

DB 181 AQAQESLDVLLRHPGSMWLVENEDQTAFPPAPYLEEAAPGQREGPSLGSSGPFQFCAS 240

QY 241 RAYESSRADELSPAGARVRVLETSDRGWMLCRYGDRAGLLPAVLLRPEGLGALLSGTGF 300

DB 241 RAYESSRADELSPAGARVRVLETSDRGWMLCRYGDRAGLLPAVLLRPEGLGALLSGTGF 300

QY 301 RCGDDPAGEARGFPFESQATAPPTVTPRSPGAIQSRCTVTRRALERRRRPQGRPGC 360

DB 301 RCGDDPAGEARGFPFESQATAPPTVTPRSPGAIQSRCTVTRRALERRRRPQGRPGC 360

QY 361 VDSVPHPTTEQ 371

DB 361 VDSVPHPTTEQ 371

RESULT 2
 ADJ57073

ID ADJ57073 standard; protein; 376 AA.

AC ADJ57073;

DT 06-MAY-2004 (first entry)

DE Nox enzyme regulatory protein p41Nox.

XX Nox regulatory protein; p41Nox; NADH oxidase; cytostatic; human.

XX Homo sapiens.

XX WO2004007689-A2.

XX 22-JAN-2004.

XX 16-JUL-2003; 2003WO-US022246.

XX 16-JUL-2002; 2002US-0396170P.

XX 23-AUG-2002; 2002US-0405647P.

XX (UYEM-) UNIV EMORY.

XX Lambeth JD, Cheng G;

XX WPI; 2004-122927/12.

XX N-PSDB; ADJ57072.

XX New regulatory protein for hydrogenated nicotinamide adenine dinucleotide

PT phosphate oxidase, useful for preparing a medicament for affecting
 PT superoxide formation or regulating cellular proliferation in an animal or
 PT a human.

PS Claim 1; SEQ ID NO 6; 70pp; English.

XX The present sequence is the protein sequence of a human regulatory
 CC protein, denoted p41Nox, for Nox (NADH oxidase) enzymes. The protein and
 CC nucleotide sequences of four p41Nox variants have been identified
 CC ADJ57068-ADJ57075. p41Nox proteins have a mol.wt. of 41 kDa. They include
 CC a PX domain and 2 SH3 domains and show sequence homology (25% identity)
 CC with p47phox, a regulatory protein for gp91phox. They function in the
 CC regulation of cell growth and are therefore implicated in diseases
 CC involving abnormal cell growth, such as cancer. They may also function in
 CC innate immune mechanisms of epithelial tissue or other barrier cells, and
 CC hence may be involved in diseases of diminished ability to fight
 CC infections or inflammatory conditions. p41Nox regulatory proteins and
 CC nucleic acids are useful for preparing a medicament for affecting
 CC superoxide formation or regulating cellular proliferation in an animal or
 CC a human. The regulatory proteins are also useful in drug development,
 CC e.g. screening for drugs that regulate the biological activity of the
 CC proteins, or in assays that relate to assessment of abnormal growth or
 CC cellular proliferation including cancer.

XX Sequence 376 AA;

Query Match 99.4%; Score 1937.5; DB 8; Length 376;
 Best Local Similarity 98.7%; Pred. No. 5.7e-167;
 Matches 371; Conservative 0; Mismatches 0; Indels 5; Gaps 1;

QY 1 MAGPRYPVSVOGAALVQIKRLQTFAPSVRWSGSDTFVRRSWDEFRQLKTKLKEFPVEA 60

DB 1 MAGPRYPVSVOGAALVQIKRLQTFAPSVRWSGSDTFVRRSWDEFRQLKTKLKEFPVEA 60

QY 61 GLLRRSDRVLPKLLDAPLLGRVGRSGRLARLQLLETYSRRLATAERVARSPIT 115

DB 61 GLLRRSDRVLPKLLDAPLLGRVGRSGRLARLQLLETYSRRLATAERVARSPIT 120

QY 116 TGFAPQPLDLEPALPGSRVILPTPEEQPLSRAAGRLSIHSLEAQSLRCLQPFCTQDTR 175

DB 121 TGFAPQPLDLEPALPGSRVILPTPEEQPLSRAAGRLSIHSLEAQSLRCLQPFCTQDTR 180

QY 176 DRPFOAQAESLDVLLRHPGSMWLVENEDQTAFPPAPYLEEAAPGQREGPSLGSSGP 235

DB 181 DRPFOAQAESLDVLLRHPGSMWLVENEDQTAFPPAPYLEEAAPGQREGPSLGSSGP 240

QY 236 QFCASRAYESSRADELSPAGARVRVLETSDRGWMLCRYGDRAGLLPAVLLRPEGLGALL 295

DB 241 QFCASRAYESSRADELSPAGARVRVLETSDRGWMLCRYGDRAGLLPAVLLRPEGLGALL 300

QY 296 SGTGFRGGDDPAGEARGFPFESQATAPPTVTPRSPGAIQSRCTVTRRALERRRRPQGR 355

DB 301 SGTGFRGGDDPAGEARGFPFESQATAPPTVTPRSPGAIQSRCTVTRRALERRRRPQGR 360

QY 356 RPRGCVDSVPHPPTTEQ 371

DB 361 RPRGCVDSVPHPPTTEQ 376

RESULT 3

ADJ57069

ID ADJ57069 standard; protein; 370 AA.

XX ADJ57069;

XX 06-MAY-2004 (first entry)

XX Nox enzyme regulatory protein p41Nox.

XX Nox regulatory protein; p41Nox; NADH oxidase; cytostatic; human.

XX Homo sapiens.

PN W02004007689-A2.
 XX 22-JAN-2004.
 XX 16-JUL-2003; 2003WO-US022246.
 XX 16-JUL-2002; 2002US-0396170P.
 PR 23-AUG-2002; 2002US-0405647P.
 XX (UYEM-) UNIV EMORY.
 XX Lambeth JD, Cheng G;
 XX WPI; 2004-122927/12.
 DR N-PSDB; ADJ57068.
 XX New regulatory protein for hydrogenated nicotinamide adenine dinucleotide
 PT phosphate oxidase, useful for preparing a medicament for affecting
 PT superoxide formation or regulating cellular proliferation in an animal or
 PT a human.
 XX
 XX Claim 1; SEQ ID NO 2; 70pp; English.
 XX
 CC The present sequence is the protein sequence of a human regulatory
 CC protein, denoted p41Nox, for Nox (NADH oxidase) enzymes. The protein and
 CC nucleotide sequences of four p41Nox variants have been identified
 CC ADJ57068-ADJ57075. p41Nox proteins have a mol.wt. of 41 kDa. They include
 CC a PX domain and 2 SH3 domains and show sequence homology (25% identity)
 CC with p47phox, a regulatory protein for gp91phox. They function in the
 CC regulation of cell growth and are therefore implicated in diseases
 CC involving abnormal cell growth, such as cancer. They may also function in
 CC innate immune mechanisms of epithelial tissue or other barrier cells, and
 CC hence may be involved in diseases of diminished ability to fight
 CC infections or inflammatory conditions. p41Nox regulatory proteins and
 CC nucleic acids are useful for preparing a medicament for affecting
 CC superoxide formation or regulating cellular proliferation in an animal or
 CC a human. The regulatory proteins are also useful in drug development,
 CC e.g. screening for drugs that regulate the biological activity of the
 CC proteins, or in assays that relate to assessment of abnormal growth or
 CC cellular proliferation including cancer.
 XX
 XX Sequence 370 AA;
 Query Match 99.2%; Score 1934.5; DB 8; Length 370;
 Best Local Similarity 99.7%; Pred. No. 1e-166;
 Matches 370; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
 QY 1 MAGPRYPVSVQGAALVQIKRLQTFAPSVRWSGSDTFVRRSWDFRQLKTLKETPVEA 60
 DB 1 MAGPRYPVSVQGAALVQIKRLQTFAPSVRWSGSDTFVRRSWDFRQLKTLKETPVEA 59
 QY 61 GLLRRSDRVLPKLLDAPLLGRVGRGRTSGRLARLQLLETYSRRLATAEVARSPPTIGFFA 120
 DB 60 GLLRRSDRVLPKLLDAPLLGRVGRGRTSGRLARLQLLETYSRRLATAEVARSPPTIGFFA 119
 QY 121 POPLDLEPALPPGSRVILPTPEEQPLSRAAGRLSIHSLAQSLRCLQPFCTQDTRDRPFQ 180
 DB 120 POPLDLEPALPPGSRVILPTPEEQPLSRAAGRLSIHSLAQSLRCLQPFCTQDTRDRPFQ 179
 QY 181 AQAESLDVLLRHPSGWLNVENEDROTAWFPAPYLEAAPGQREGGPGSLGSGGPOFCAS 240
 DB 180 AQAESLDVLLRHPSGWLNVENEDROTAWFPAPYLEAAPGQREGGPGSLGSGGPOFCAS 239
 QY 241 RAYESSRADELSVPAGARVRLTSDRGWMLCRYGDRAGLPAVLRLPEGLGALLSGTGF 300
 DB 240 RAYESSRADELSVPAGARVRLTSDRGWMLCRYGDRAGLPAVLRLPEGLGALLSGTGF 299
 QY 301 RGGDDPAGEARGPPEPSQATAPPTVTPRESGAIQSRCCCTVTRRALERRPRQGRPGC 360
 DB 300 RGGDDPAGEARGPPEPSQATAPPTVTPRESGAIQSRCCCTVTRRALERRPRQGRPGC 359
 QY 361 VDSVPHPTTEQ 371
 |||||||

Db 360 VDSVPHPTTEQ 370
 RESULT 4
 ADJ57075
 ID ADJ57075 standard; protein; 375 AA.
 XX AC ADJ57075;
 XX 06-MAY-2004 (first entry)
 XX Nox enzyme regulatory protein p41Nox.
 DE Nox regulatory protein; p41Nox; NADH oxidase; cytostatic; human.
 XX
 XX Homo sapiens.
 PN W02004007689-A2.
 XX 22-JAN-2004.
 XX 16-JUL-2003; 2003WO-US022246.
 XX 16-JUL-2002; 2002US-0396170P.
 PR 23-AUG-2002; 2002US-0405647P.
 XX (UYEM-) UNIV EMORY.
 XX Lambeth JD, Cheng G;
 XX WPI; 2004-122927/12.
 DR N-PSDB; ADJ57074.
 XX New regulatory protein for hydrogenated nicotinamide adenine dinucleotide
 PT phosphate oxidase, useful for preparing a medicament for affecting
 PT superoxide formation or regulating cellular proliferation in an animal or
 PT a human.
 XX
 XX Claim 1; SEQ ID NO 8; 70pp; English.
 XX
 CC The present sequence is the protein sequence of a human regulatory
 CC protein, denoted p41Nox, for Nox (NADH oxidase) enzymes. The protein and
 CC nucleotide sequences of four p41Nox variants have been identified
 CC ADJ57068-ADJ57075. p41Nox proteins have a mol.wt. of 41 kDa. They include
 CC a PX domain and 2 SH3 domains and show sequence homology (25% identity)
 CC with p47phox, a regulatory protein for gp91phox. They function in the
 CC regulation of cell growth and are therefore implicated in diseases
 CC involving abnormal cell growth, such as cancer. They may also function in
 CC innate immune mechanisms of epithelial tissue or other barrier cells, and
 CC hence may be involved in diseases of diminished ability to fight
 CC infections or inflammatory conditions. p41Nox regulatory proteins and
 CC nucleic acids are useful for preparing a medicament for affecting
 CC superoxide formation or regulating cellular proliferation in an animal or
 CC a human. The regulatory proteins are also useful in drug development,
 CC e.g. screening for drugs that regulate the biological activity of the
 CC proteins, or in assays that relate to assessment of abnormal growth or
 CC cellular proliferation including cancer.
 XX
 XX Sequence 375 AA;
 Query Match 98.6%; Score 1922; DB 8; Length 375;
 Best Local Similarity 98.4%; Pred. No. 1.4e-165;
 Matches 370; Conservative 0; Mismatches 0; Indels 6; Gaps 2;
 QY 1 MAGPRYPVSVQGAALVQIKRLQTFAPSVRWSGSDTFVRRSWDFRQLKTLKETPVEA 60
 DB 1 MAGPRYPVSVQGAALVQIKRLQTFAPSVRWSGSDTFVRRSWDFRQLKTLKETPVEA 59
 QY 61 GLLRRSDRVLPKLLDAPLLGRVGRGRTSGRLARLQLLETYSRRLATAEVARSPPTI 115
 DB 60 GLLRRSDRVLPKLLDAPLLGRVGRGRTSGRLARLQLLETYSRRLATAEVARSPPTI 119
 QY 116 TGFFAPQPLDLEPALPPGSRVILPTPEEQPLSRAAGRLSIHSLAQSLRCLQPFCTQDTR 175
 |||||||

Db 120 TGFFAPQPLDLBPALPPGSRVILPTPEEQPLSRAAGRLSIHSLEAQSRLCLQPFCTQDTR 179
QY 176 DRPFOAQAESLDVLLRHPSGWLWENEDROTAMFPAPYLEEAPQGREGGSLGSSGP 235
Db 180 DRPFOAQAESLDVLLRHPSGWLWENEDROTAMFPAPYLEEAPQGREGGSLGSSGP 239
QY 236 QFCASRAYESSRADELSVPAGARVRVLETSRGMWLCRYGDRAGLLPAVILRPEGIGALL 295
Db 240 QFCASRAYESSRADELSVPAGARVRVLETSRGMWLCRYGDRAGLLPAVILRPEGIGALL 299
QY 296 SGTGFRGDDPAGEARGFPEPSQATAPPTVTRPSGAIQSRCCCTVTRALLERPRRQG 355
Db 300 SGTGFRGDDPAGEARGFPEPSQATAPPTVTRPSGAIQSRCCCTVTRALLERPRRQG 359
QY 356 RPRGCVDSVPHTTEQ 371
Db 360 RPRGCVDSVPHTTEQ 375

RESULT 5
ABG24891
ID ABG24891 standard; protein; 419 AA.
XX AC ABG24891;
XX DT 18-FEB-2002 (first entry)
XX DE Novel human diagnostic protein #24882.

XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX KW food supplement; medical imaging; diagnostic; genetic disorder.
XX OS Homo sapiens.
XX PN WO200175067-A2.
XX PD 11-OCT-2001.

XX PF 30-MAR-2001; 2001WO-US008631.
XX PR 31-MAR-2000; 2000US-00540217.
XX PR 23-AUG-2000; 2000US-00649167.
XX PA (HYSE-) HYSEQ INC.
XX PI Drmanac RT, Liu C, Tang YT;
XX DR WPI; 2001-639362/73.
XX DR N-PSDB; AAS89078.

XX PT New isolated polynucleotide and encoded polypeptides, useful in
XX PT diagnostics, forensics, gene mapping, identification of mutations
XX PT responsible for genetic disorders or other traits and to assess
XX PT biodiversity.
XX PS Claim 20; SEQ ID NO 55250; 103pp; English.

XX CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
XX CC sequences. (I) is useful as hybridisation probes, polymerase chain
XX CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
XX CC and in recombinant production of (II). The polynucleotides are also used
XX CC in diagnostics as expressed sequence tags for identifying expressed
XX CC genes. (I) is useful in gene therapy techniques to restore normal
XX CC activity of (II) or to treat disease states involving (II). (II) is
XX CC useful for generating antibodies against it, detecting or quantitating a
XX CC polypeptide in tissue, as molecular weight markers and as a food
XX CC supplement. (II) and its binding partners are useful in medical imaging
XX CC of sites expressing (II). (I) and (II) are useful for treating disorders
XX CC involving aberrant protein expression or biological activity. The
XX CC polypeptide and polynucleotide sequences have applications in
XX CC diagnostics, forensics, gene mapping, identification of mutations
XX CC responsible for genetic disorders or other traits to assess biodiversity

CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
CC amino acid sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 419 AA;

Query Match 72.7%; Score 1418; DB 4; Length 419;
Best Local Similarity 95.9%; Pred. No. 9.1e-120;
Matches 278; Conservative 0; Mismatches 4; Indels 8; Gaps 2;

QY 1 MAGPRYPVSVOGAAALVQIKRLQTFAPSVRWSDGSDTFVRRSWDEPRLKTKLKEFPVEA 60
Db 1 MAGPRYPVSVOGAAALVQIKRLQTFAPSVRWSDGSDTFVRRSWDEPRLKTKLKEFPVEA 60
QY 61 GLLRRSDRVLPKL-----LDAPLLGRVGTSGRLARLQLLETYSRRLATARVARSPIT 115
Db 61 GLLRRSDRVLPKLGLQASLDAPLLGRVGTSGRLARLQLLETYSRRLATARVARSPIT 120
QY 116 TGFFAPQPLDLBPALPPGSRVILPTPEEQPLSRAAGRLSIHSLEAQSRLCLQPFCTQDTR 175
Db 121 TGFFAPQPLDLBPALPPGSRVILPTPEEQPLSRAAGRLSIHSLEAQSRLCLQPFCTQDTR 180
QY 176 DRPFOAQAESLDVLLRHPSGWLWENEDROTAMFPAPYLEEAPQGREGGSLGSSGP 235
Db 181 DRPFOAQAESLDVLLRHPSGWLWENEDROTAMFPAPYLEEAPQGREGGSLGSSGP 240
QY 236 QFCASRAYESSRADELSVPAGARVRVLETSRGMWLCRYGDRAGLLPAVL 285
Db 241 QFCASRAYESSRADELSVPAGARVRVLETSRGMWLCRY---AGAGPEEL 287

RESULT 6
ABG25247
ID ABG25247 standard; protein; 726 AA.
XX AC ABG25247;
XX DT 18-FEB-2002 (first entry)
XX DE Novel human diagnostic protein #25238.
XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX KW food supplement; medical imaging; diagnostic; genetic disorder.
XX OS Homo sapiens.
XX PN WO200175067-A2.
XX PD 11-OCT-2001.
XX PF 30-MAR-2001; 2001WO-US008631.
XX PR 31-MAR-2000; 2000US-00540217.
XX PR 23-AUG-2000; 2000US-00649167.
XX PA (HYSE-) HYSEQ INC.
XX PI Drmanac RT, Liu C, Tang YT;
XX DR WPI; 2001-639362/73.
XX DR N-PSDB; AAS89434.
XX PT New isolated polynucleotide and encoded polypeptides, useful in
XX PT diagnostics, forensics, gene mapping, identification of mutations
XX PT responsible for genetic disorders or other traits and to assess
XX PT biodiversity.
XX PS Claim 20; SEQ ID NO 55606; 103pp; English.
XX CC The invention relates to isolated polynucleotide (I) and polypeptide (II)

sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activities. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic amino acid sequences of the invention. Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX Sequence 726 AA;

Query Match 72.7%; Score 1418; DB 4; Length 726;
Best Local Similarity 95.9%; Pred. No. 1.9e-119;
Matches 278; Conservative 0; Mismatches 4; Indels 8; Gaps 2;

QY 1 MAGPRYPVSVOGAALVQIKRLQTFAPSVRWSDGSDTFVRRSWDFRQLKTKLTPVEA 60
DB 1 MAGPRYPVSVOGAALVQIKRLQTFAPSVRWSDGSDTFVRRSWDFRQLKTKLTPVEA 60

QY 61 GLLRSDRVLPKL-----LDAPLLGRVGRTRGLARLQLLETYSRLLATAERVARSPTI 115
DB 61 GLLRSDRVLPKLQASLDAPLLGRVGRTRGLARLQLLETYSRLLATAERVARSPTI 120

QY 116 TGFFAPQPLDLEPALPPGSRVILPTPEOPLSRAAGRLSIHSLEAQSRLCLQPFCTQDTR 175
DB 121 TGFFAPQPLDLEPALPPGSRVILPTPEOPLSRAAGRLSIHSLEAQSRLCLQPFCTQDTR 180

QY 176 DRPFOAQAESLDVLLRHPGSMWLVNEDRQTAWFPAPYLEEAAPQGGREGGSLGSSGP 235
DB 181 DRPFOAQAESLDVLLRHPGSMWLVNEDRQTAWFPAPYLEEAAPQGGREGGSLGSSGP 240

QY 236 QFCASRAYESSRADELSPAGARVRVLETSDRGWMLCRYGDRAGLLPAVL 285
DB 241 QFCASRAYESSRADELSPAGARVRVLETSDRGWMLCRY---AGAGPEEL 287

RESULT 7
ADG36779
ID ADG36779 standard; protein; 390 AA.
XX
AC ADG36779;
XX
XX 26-FEB-2004 (first entry)
XX
DE Human P47PHOX-related protein.
XX
KW human; autoimmune condition; NADPH oxidase activity;
KW NADPH oxidase activator; P47PHOX; arthritis; multiple sclerosis; lupus;
KW autoimmune uveitis; type I diabetes; bronchial asthma; septic arthritis;
KW cardiovascular disease.
OS Homo sapiens.
XX
XX WO2003095667-A2.
XX
XX 20-NOV-2003.
XX
XX 13-MAY-2003; 2003WO-IB002419.
XX
XX 13-MAY-2002; 2002US-0380904P.
PR
PR 27-NOV-2002; 2002US-0429609P.

XX (AREX-) AREXIS AB.
XX Holmdahl R, Olofsson P;
XX WPI; 2004-012133/01.
DR
XX Assessing a mammal's susceptibility to develop an autoimmune condition by
PT determining whether or not the level of NADPH oxidase activity of the
PT cell after contacting the cell with an NADPH oxidase activator is less
PT than a control level.
XX
PS Disclosure; SEQ ID NO 7; 103pp; English.
XX
CC The invention comprises a method for assessing the susceptibility of a
CC mammal to develop an autoimmune condition. The method involves: providing
CC a blood or synovial fluid sample containing a cell from a mammal
CC determining the level of NADPH oxidase activity of the cell after
CC contacting the cell with an NADPH oxidase activator, and determining
CC whether or not the level is less than a control level of NADPH oxidase
CC activity. The method further comprises determining whether or not a
CC mammal has a genetic variant of the gene encoding a polypeptide that
CC functions in the NADPH oxidase pathway (e.g. P47PHOX), where the presence
CC of the genetic variant indicates that the mammal is susceptible to
CC develop an autoimmune condition. The method is useful in the diagnosis
CC and treatment of autoimmune conditions, such as: arthritis, multiple
CC sclerosis, lupus, autoimmune uveitis, type I diabetes, bronchial asthma,
CC septic arthritis and cardiovascular disease. The present amino acid
CC sequence represents a human P47PHOX-related protein.

XX Sequence 390 AA;

Query Match 15.9%; Score 309.5; DB 8; Length 390;
Best Local Similarity 25.5%; Pred. No. 3.2e-19;
Matches 96; Conservative 62; Mismatches 151; Indels 67; Gaps 13;

QY 24 FAFSVRWSDGSDTFVRRSWDFRQLKTKLTFPVEAGLLRSDRVLPKLLDAPLLGRVG 83
DB 26 YMFVLKWDLSKVVVRRFTETVEFHKTKEMFPIEAGAINPENRIIPH-LPAPKWFDDG 84

QY 84 RTSRGLARLQLLETYSRLLATAERVARSPITITGFAPQPLDLEPALPPGSRV-----I 137
DB 85 RAAEN--RQGLTEYCSLMSLPTKISRCPHLLDFKVRPDDLK--LPTDNTQKKPETYL 140

QY 138 LPTPEOPLSRAAGRLSIHSLEAQSRLCLQPFCTQDTRDRPFOAQAESLDVLLRHPGSM 197
DB 141 MPDQKSTATDITGPILL-----QSYRAIANY--EKTSGSENALSTGCVVEVVEKESGW 193

QY 198 WLVENEDRQTAWFPAPYLEEA-APQGGREGGSLGSSGPFQFCASRAYESSRADELSPVAG 256
DB 194 WFCQMKAKR-GWIPASFLPLELSDPDETEDEPNY--AGEPYVAIKAYTAVEGDEVSLLEG 250

QY 257 ARVRVLETSDRGWMLCRYGDRAGLLPAVLLRREG-----LGA----- 293
DB 251 EAVEVTHKLLDGMWVIRKDDVTGYFFSMYLOKSGQDVSOAQRIKRGAPPRSRINVHS 310

QY 294 -----LLSGTGFRRGGD-----DPAGEARGEP-----EPSQATAPPPTVTPRS 331
DB 311 IHQRGRKRLSQDAYRRNSVRFLQRRRRQARPGFPGSPLEERQTQRKSPQPAVPPRS 370

QY 332 PGAIQSRCCCTVTRRAL 347

DB 371 ADLILNRCSESTYRKL 386

RESULT 8
ADG36778
ID ADG36778 standard; protein; 390 AA.
XX
XX ADG36778;
AC
XX
XX 26-FEB-2004 (first entry)
DT
XX

DE Human P47PHOX protein #2.
XX human; autoimmune condition; NADPH oxidase activity;
KW NADPH oxidase activator; P47PHOX; arthritis; multiple sclerosis; lupus;
KW autoimmune uveitis; type I diabetes; bronchial asthma; septic arthritis;
KW cardiovascular disease.
XX
XX
OS Homo sapiens.
XX
FN WO2003095667-A2.
XX
XX 20-NOV-2003.
XX
XX 13-MAY-2003; 2003WO-IB002419.
XX
PR 13-MAY-2002; 2002US-0380904P.
PR 27-NOV-2002; 2002US-0429609P.
XX
XX (AREX-) AREXIS AB.
XX
XX Holmdahl R, Olofsson P;
XX
DR WPI; 2004-012133/01.
XX N-PSDB; ADG36777.
XX
PT Assessing a mammal's susceptibility to develop an autoimmune condition by
PT determining whether or not the level of NADPH oxidase activity of the
PT cell after contacting the cell with an NADPH oxidase activator is less
PT than a control level.
XX
XX Claim 17; SEQ ID NO 6; 103pp; English.
XX
XX The invention comprises a method for assessing the susceptibility of a
CC mammal to develop an autoimmune condition. The method involves: providing
CC a blood or synovial fluid sample containing a cell from a mammal,
CC determining the level of NADPH oxidase activity of the cell after
CC contacting the cell with an NADPH oxidase activator, and determining
CC whether or not the level is less than a control level of NADPH oxidase
CC activity. The method further comprises determining whether or not a
CC mammal has a genetic variant of the gene encoding a polypeptide that
CC functions in the NADPH oxidase pathway (e.g. P47PHOX), where the presence
CC of the genetic variant indicates that the mammal is susceptible to
CC develop an autoimmune condition. The method is useful in the diagnosis
CC and treatment of autoimmune conditions, such as: arthritis, multiple
CC sclerosis, lupus, autoimmune uveitis, type I diabetes, bronchial asthma,
CC septic arthritis and cardiovascular disease. The present amino acid
CC sequence represents a human P47PHOX protein.
XX
XX Sequence 390 AA;
XX
XX Query Match 15.8%; Score 307.5; DB 8; Length 390;
XX Best Local Similarity 25.3%; Pred. No. 4.9e-19;
XX Matches 95; Conservative 63; Mismatches 151; Indels 67; Gaps 13;
XX
QY 24 FAFSVRWSDGSDTFFRRSWDEFRQLKTKLTFPVEAGLLRRSDRVLPKLLDAPLLGRVG 83
DB 26 YMFVLKQDLSEKVVYRRFTIYEFHKLTFEMFFIEAGINPENRIIPH-LPAKWFQDG 84
XX
QY 84 RTSRGLARLQLLETYSRRLATAERVARSPITITGFFAPQPLDLEPALPPGSRV-----I 137
DB 85 RAAEN--RQGTLTTCSTLMSLPTKISRCPHLLDFKVRPDDLK--LPTDQTKKPEYTL 140
XX
QY 138 LPTPEEQLSRAGRLSIHSLEAQSLRCLQPFCTQDTRDRPFOAQASLDVLLRHPSGW 197
DB 141 MPKDGKSTATDITGPIIL-----QTYRAIDY--EKTSGSEMALSTGVDVVEVKESEGW 193
XX
QY 198 WLVENEDROTAFWPAPYLEEA-APCGREGGPGSLGGSPQFCASRAYESSRADELSVPAG 256
DB 194 WFCQMKAR-GWIPASFLEPLDSPDETDPENY--AGEPYAIKAYTAVGEVDSVLEGG 250
XX
QY 257 ARVRVLETSDRGWMLCRYDGRAGLLPAVLLRPEG-----LGA----- 293
DB 251 EAVEVIHLLDGMVIRKDDVTGYFFSMYLOKSGQDVSQAQRQIKRGAPRRSRIRNAHS 310
XX

QY 294 -----LLSCTGPRGGD-----DPAGEARGFP-----EPQATADPTTVPTRPS 331
DB 311 IHQSRRLQLQDAYRRNSVRFLOQRRRQARPGSPGSPLEERQTSKSKQPAVPFRPS 370
QY 332 PGAIQSRCCCTVTRRAL 347
DB 371 ADLILNRCSESTKTKL 386
XX
XX RESULT 9
XX ADG62962
ID ADG62962 standard; protein; 390 AA.
XX AC ADG62962;
XX
XX 11-MAR-2004 (first entry)
XX
XX p47(phox).
XX
XX NAD(P)H oxidase; inhibitor; blood glucose; pyridine; imidazole;
KW diethyl pyrocarbonate; chloromercuribenzoic acid;
KW 4-(2-aminomethyl)sulfonyl fluoride acetovanillone; type II; diabetes.
XX
XX Homo sapiens.
XX
XX WO2003087399-A1.
XX
XX 23-OCT-2003.
XX
XX 16-APR-2003; 2003WO-SE000618.
XX
XX 17-APR-2002; 2002SE-00001152.
XX 13-SEP-2002; 2002US-0410626P.
XX
XX (BIOV-) BIOVITRUM AB.
XX
XX Groenberg A, Wikstroem P;
XX WPI; 2004-081851/08.
XX N-PSDB; ADG62961.
XX
XX Treatment of a medical condition characterized by elevated blood glucose
XX levels, especially diabetes, comprises administering a NAD(P)H oxidase
XX inhibitor.
XX
XX Disclosure; SEQ ID NO 18; 89pp; English.
XX
XX This sequence represents the NAD(P)H oxidase, p47(phox). This protein may
XX be used to identify a NAD(P)H oxidase inhibitor which may be administered
XX to a patient in the method of the invention for treatment of a medical
XX condition characterized by elevated blood glucose levels. The NAD(P)H
XX oxidase inhibitor is preferably pyridine, imidazole, diethyl
XX pyrocarbonate, chloromercuribenzoic acid or a substance referred to as 4-
XX (2-aminomethyl)sulfonyl fluoride acetovanillone. The NAD(P)H oxidase
XX inhibitor may be used for treating diabetes, especially type II diabetes.
XX
XX Sequence 390 AA;
XX
XX Query Match 15.8%; Score 307.5; DB 8; Length 390;
XX Best Local Similarity 25.3%; Pred. No. 4.9e-19;
XX Matches 95; Conservative 63; Mismatches 151; Indels 67; Gaps 13;
XX
QY 24 FAFSVRWSDGSDTFFRRSWDEFRQLKTKLTFPVEAGLLRRSDRVLPKLLDAPLLGRVG 83
DB 26 YMFVLKQDLSEKVVYRRFTIYEFHKLTFEMFFIEAGINPENRIIPH-LPAKWFQDG 84
XX
QY 84 RTSRGLARLQLLETYSRRLATAERVARSPITITGFFAPQPLDLEPALPPGSRV-----I 137
DB 85 RAAEN--RQGTLTTCSTLMSLPTKISRCPHLLDFKVRPDDLK--LPTDQTKKPEYTL 140
XX
QY 138 LPTPEEQLSRAGRLSIHSLEAQSLRCLQPFCTQDTRDRPFOAQASLDVLLRHPSGW 197
DB 251 EAVEVIHLLDGMVIRKDDVTGYFFSMYLOKSGQDVSQAQRQIKRGAPRRSRIRNAHS 310
XX

Db 141 MPKDGKSTATDITGPIL-----QTVRAIADY--EKTSGSEMALSTGDDVVVVEKSESGW 193
 QY 198 WLVENEDROTAMFPAPYLEEA-APGQREGGSLGSSGPOFCASRAYESSRADELSPAG 256
 Db 194 WFCQMKAKR-GWIPASFLPLDSPDETEDEPENY--AGEPYVAIKAYTAVEGDEVSLLEG 250
 QY 257 ARVRVLETSRQWNLCRYGDRAGLLPAVLRLRPEG-----LGA----- 293
 Db 251 EAVEVIHKLDDGWWVIRKDDVTGYFPMYLOKSGQDVSAQROIKRGAPPRSSIRNAHS 310
 QY 294 -----LLSGTGFRGDD-----DPAGEARGFP-----EPSQATAPPTVTPRPS 331
 Db 311 IHQSRKRLSQDAYRRNSVRFLOQRRRQARPGQSPGSPLEERQTRQSKPQPAVPPRPS 370
 QY 332 PGAIQSRCCCTVTRRAL 347
 Db 371 ADLILNRCSESTKRL 386

RESULT 10
 ID ADP23795 standard; protein; 390 AA.
 XX
 AC ADP23795;
 XX
 DT 18-NOV-2004 (first entry)
 XX
 DE PRO polypeptide SEQ ID NO:973.
 DE
 KW PRO; antiinflammatory; antiarthritic; antirheumatic; immunosuppressive;
 KW osteopathic; antidiabetic; dermatological; antipsoriatic; antiallergic;
 KW antiasthmatic; hepatotropic; respiratory; gene therapy; immune system.
 OS Unidentified.
 XX
 PN WO2004041170-A2.
 XX
 PD 21-MAY-2004.
 XX
 PF 30-OCT-2003; 2003WO-US034312.
 XX
 PR 01-NOV-2002; 2002US-0423394P.
 XX
 PA (GETH) GENENTECH INC.
 PI Clark H, Schoenfeld J, Van Lookeren M, Williams PM, Wood WI;
 PI Wu TD;
 DR WPI; 2004-419628/39.
 DR N-PSDB; ADP23794.
 XX
 XX New PRO polypeptides and polynucleotides, useful for treating e.g.
 PT erythematous, rheumatoid arthritis, diabetes mellitus, immune-mediated
 PT renal disease, or demyelinating diseases of the central or peripheral
 PT nervous system.
 XX
 PS Claim 7; SEQ ID NO 973; 2940pp; English.
 XX
 CC The invention relates to a novel isolated nucleic acid and the PRO
 CC polypeptide encoded by it. A protein of the invention has
 CC antiinflammatory, antiarthritic, antirheumatic, immunosuppressive,
 CC osteopathic, antidiabetic, dermatological, antipsoriatic, antiallergic,
 CC antiasthmatic, hepatotropic, and respiratory activity. A polynucleotide
 CC of the invention may have a use in gene therapy. The PRO polypeptide, its
 CC agonist, antagonist, or antibody that specifically binds to the
 CC polypeptide is useful for treating an immune related disorder such as
 CC juvenile lupus erythematous, rheumatoid arthritis, osteoarthritis,
 CC systemic chronic arthritis, a spondyloarthropathy, systemic sclerosis, an
 CC idiopathic inflammatory myopathy, Sjogren's syndrome, systemic
 CC vasculitis, sarcoidosis, autoimmune haemolytic anaemia, autoimmune
 CC thrombocytopenia, thyroiditis, diabetes mellitus, immune-mediated renal
 CC disease, a demyelinating disease of the central or peripheral nervous
 CC system, idiopathic demyelinating polyneuropathy, Guillain-Barre syndrome,

CC a chronic inflammatory demyelinating polyneuropathy, a hepatobiliary
 CC disease, infectious or autoimmune chronic active hepatitis, primary
 CC biliary cirrhosis, granulomatous hepatitis, sclerosing cholangitis,
 CC inflammatory bowel disease, gluten-sensitive enteropathy, Whipple's
 CC disease, an autoimmune or immune-mediated skin disease, a bullous skin
 CC disease, erythema multiforme, contact dermatitis, psoriasis, an allergic
 CC disease, asthma, allergic rhinitis, atopic dermatitis, food
 CC hypersensitivity, urticaria, an immunologic disease of the lung,
 CC eosinophilic pneumonia, idiopathic pulmonary fibrosis, hypersensitivity
 CC pneumonitis, a transplantation associated disease, graft rejection or
 CC graft-versus-host disease. The present sequence represents a PRO protein
 CC of the invention.
 XX
 SQ Sequence 390 AA;
 Query Match 15.8%; Score 307.5; DB 8; Length 390;
 Best Local Similarity 25.3%; Pred. No. 4.9e-19;
 Matches 95; Conservative 63; Mismatches 151; Indels 67; Gaps 13;
 QY 24 FAFSVRWSDGSDTFVRRSWDEFRQLKTKLTFPVEAGLLRRSDRVLPKLLDAPLLGRVG 83
 Db 26 YMFVKWQDLSEKVVVRRFTETVEYFHKTLKEMFPIEAGAINPENRIIPH-LPAPKWFQDQ 84
 QY 84 RTSRGLARQLLETYSRRLLATAERVARSPPTTGTFFAPQDLLEPALPGSRV-----I 137
 Db 85 RAAEN--RQGTLTETCYCSTLMSLPTKISRCPHLLDFKVRPDDLK--LPTDNOTKKPETYL 140
 QY 138 LPTPEEQPLSRAAGRLSIHSLAQSLRCLQPFCTQDTRDRPFAQAQESLSDVLLRHPSGW 197
 Db 141 MPKDGKSTATDITGPIL-----QTVRAIADY--EKTSGSEMALSTGDDVVVVEKSESGW 193
 QY 198 WLVENEDROTAMFPAPYLEEA-APGQREGGSLGSSGPOFCASRAYESSRADELSPAG 256
 Db 194 WFCQMKAKR-GWIPASFLPLDSPDETEDEPENY--AGEPYVAIKAYTAVEGDEVSLLEG 250
 QY 257 ARVRVLETSRQWNLCRYGDRAGLLPAVLRLRPEG-----LGA----- 293
 Db 251 EAVEVIHKLDDGWWVIRKDDVTGYFPMYLOKSGQDVSAQROIKRGAPPRSSIRNAHS 310
 QY 294 -----LLSGTGFRGDD-----DPAGEARGFP-----EPSQATAPPTVTPRPS 331
 Db 311 IHQSRKRLSQDAYRRNSVRFLOQRRRQARPGQSPGSPLEERQTRQSKPQPAVPPRPS 370
 QY 332 PGAIQSRCCCTVTRRAL 347
 Db 371 ADLILNRCSESTKRL 386

RESULT 11
 ADG36776
 ID ADG36776 standard; protein; 389 AA.
 XX
 AC ADG36776;
 XX
 DT 26-FEB-2004 (first entry)
 XX
 DE Human P47PHOX protein #1.
 DE
 KW human; autoimmune condition; NADPH oxidase activity;
 KW NADPH oxidase activator; P47PHOX; arthritis; multiple sclerosis; lupus;
 KW autoimmune uveitis; type I diabetes; bronchial asthma; septic arthritis;
 KW cardiovascular disease.
 OS Homo sapiens.
 XX
 PN WO2003095667-A2.
 XX
 PD 20-NOV-2003.
 XX
 PF 13-MAY-2003; 2003WO-IB002419.
 XX
 PR 13-MAY-2002; 2002US-0380904P.
 PR 27-NOV-2002; 2002US-0429609P.

```
XX (AREX-) AREXIS AB.
XX
XX Holmdahl R, Olofsson P;
XX
XX WPI; 2004-012133/01.
XX N-PSDB; ADG36775.
XX
XX Assessing a mammal's susceptibility to develop an autoimmune condition by
XX determining whether or not the level of NADPH oxidase activity of the
XX cell after contacting the cell with an NADPH oxidase activator is less
XX than a control level.
XX
XX Claim 18; SEQ ID NO 4; 103pp; English.
XX
XX The invention comprises a method for assessing the susceptibility of a
XX mammal to develop an autoimmune condition. The method involves: providing
XX a blood or synovial fluid sample containing a cell from a mammal,
XX determining the level of NADPH oxidase activity of the cell after
XX contacting the cell with an NADPH oxidase activator, and determining
XX whether or not the level is less than a control level of NADPH oxidase
XX activity. The method further comprises determining whether or not a
XX mammal has a genetic variant of the gene encoding a polypeptide that
XX functions in the NADPH oxidase pathway (e.g. P47PHOX), where the presence
XX of the genetic variant indicates that the mammal is susceptible to
XX develop an autoimmune condition. The method is useful in the diagnosis
XX and treatment of autoimmune conditions, such as: arthritis, multiple
XX sclerosis, lupus, autoimmune uveitis, type I diabetes, bronchial asthma,
XX septic arthritis and cardiovascular disease. The present amino acid
XX sequence represents a human P47PHOX protein.
XX
XX Sequence 389 AA;
XX
XX Query Match 15.7%; Score 307; DB 8; Length 389;
XX Best Local Similarity 27.1%; Pred. No. 5.4e-19;
XX Matches 101; Conservative 55; Mismatches 155; Indels 62; Gaps 14;
XX
XX QY 24 FAFSVRWSDGDTFVRSWDEPRLKTKLKEFPVEAGLLRRSDRVLPKLLDAPLLGRVG 83
XX Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
XX 26 YNFLVKQDLSKVVYRKFTIYERPHKMLKMFPIERGEIHTEKRVIPH-LPAPRWYDQG 84
XX QY 84 RTSRGLARQLLETYSRELLATAERVARSPITTFPAPQLDLEPALPPGSRVILPTPEE 143
XX Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
XX 85 RAAE--SRQGITLTFVNSLMGLPVKISRCPHLLNFKVRPDDLK--LNDQVKKP---E 137
XX QY 144 QLSRAAGRLSIHSLAE---QSLRCLQFFCTQDTRDRPFOAQAESLDVLLRHPSGWL 199
XX Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
XX 138 TYLTAKDGKNNVADITGPILLQTYRAIADY--EKSGKTEMTVATGDVVDVVEKSESGWVF 195
XX QY 200 VENEDROTAWPAPYLEEA-APQGREGPSLGGSGPQFCASRAYESSRADELSVPAGAR 258
XX Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
XX 196 COMTKR-GWVPASLYELDPSDEADDPDNY--AGEPYVTIKAYAAVEDEVSLSEGEA 252
XX QY 259 VRVLETSRQGWMLCRYGDRAGLLPAVLLRPEG-----LGA----- 293
XX Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
XX 253 LEVIHKLGDGWWVRKGIITGVFSPMYLQAGEEITQQRQIRSGAPRPRSTIRNAQSI 312
XX QY 294 -----LLSGTGFRGD-----DPAGEARGPPE-----PSQATA-PPPTVTPTRPSGA 334
XX Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
XX 313 HQSRKRLSQDITYRRNSVRFLQORRRPARPQSPDSDKDNFSTPRAKQPQAPVPRPSSDL 372
XX QY 335 IQSRCTVTRAL 347
XX Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
XX 373 ILHRCCTESTKKKL 385
XX
XX RESULT 12
XX ADRI4450
XX ID ADRI4450 standard; protein; 390 AA.
XX AC
XX ADRI4450;
XX
XX DT 21-OCT-2004 (first entry)
```

```
XX Human NF-kappaB pathway-associated protein SeqID451.
XX
XX NF-kappaB pathway; antiinflammatory; cytostatic; hepatotropic; virucide;
XX antiarthritic; antirheumatic; gastrointestinal-Gen; antiasthmatic;
XX antiarteriosclerotic; immunomodulator; cerebroprotective; vasotropic;
XX immunosuppressive; vulnary; gene therapy; immune disorder;
XX inflammatory disorder; NF-kappaB regulation; cancer; aberrant apoptosis;
XX hepatic disorder; Hodgkin's lymphoma; haematopoietic tumour;
XX hyper-IgM syndrome; hypohidrotic ectodermal dysplasia;
XX X-linked anhidrotic ectodermal dysplasia; immunodeficiency;
XX viral infection; HIV-1; HTLV-1; hepatitis B; hepatitis C; EBV; influenza;
XX rheumatoid arthritis; inflammatory bowel disease; colitis; asthma;
XX atherosclerosis; cachexia; euthyroid sick syndrome; stroke; EAE;
XX autoimmune disorder; hyper immune activity;
XX aberrant acute phase response; hypercongenital condition; birth defect;
XX necrotic lesion; wound; organ transplant rejection;
XX aberrant signal transduction; proliferating disorder; cancer;
XX HIV propagation; human.
XX
XX Homo sapiens.
XX
XX WO2004065577-A2.
XX
XX 05-AUG-2004.
XX
XX 13-JAN-2004; 2004MO-US0000798.
XX
XX 14-JAN-2003; 2003US-0440068P.
XX 12-MAY-2003; 2003US-0469757P.
XX
XX (BRIM ) BRISTOL-MYERS SQUIBB CO.
XX
XX Nadler SG, Neubauer MG, Feder JN, Carman J;
XX
XX WPI; 2004-562168/54.
XX N-PSDB; ADR14451.
XX
XX New isolated polynucleotides and polypeptides associated with NF-kappaB
XX pathway, useful for diagnosing, treating, or preventing disorders or
XX diseases associated with NF-kappaB pathway.
XX
XX Claim 1; SEQ ID NO 451; 237pp; English.
XX
XX This invention relates to the novel association of protein sequences (and
XX the genes which encode them) to the NF-kappaB pathway. The invention may
XX be useful for the production of compounds with an antiinflammatory,
XX cytostatic, hepatotropic, virucide, antiarthritic, antirheumatic,
XX gastrointestinal-Gen, antiasthmatic, antiarteriosclerotic,
XX immunomodulator, cerebroprotective, vasotropic, immunosuppressive or
XX vulnary activity or for gene therapy. The proteins and nucleotides are
XX useful for diagnosing, preventing, treating, or ameliorating conditions
XX or diseases associated with the NF-kappaB pathway. The condition is an
XX immune disorder, an inflammatory disorder, an inflammatory disorder
XX related to aberrant NF-kappaB regulation, cancer, aberrant apoptosis,
XX hepatic disorders, Hodgkin's lymphomas, haematopoietic tumours, hyper-IgM
XX syndromes, hypohidrotic ectodermal dysplasia, X-linked anhidrotic
XX ectodermal dysplasia, immunodeficiency, viral infections, HIV-1, HTLV-1,
XX hepatitis B, hepatitis C, EBV, influenza, viral replication, host cell
XX survival, evasion of immune responses, rheumatoid arthritis, inflammatory
XX bowel disease, colitis, asthma, atherosclerosis, cachexia, euthyroid sick
XX syndrome, stroke, EAE, autoimmune disorders, disorders related to hyper
XX immune activity, disorders related to aberrant acute phase responses,
XX hypercongenital conditions, birth defects, necrotic lesions, wounds,
XX organ transplant rejection, conditions related to organ transplant
XX rejection, disorders related to aberrant signal transduction,
XX proliferating disorders, cancers and HIV propagation in cells infected
XX with other viruses. The present sequence is that of a human protein which
XX is subject to the novel association with the NF-kappaB pathway of the
XX invention. Note: This sequence does not appear in the specification but
XX was obtained by the indexer from Genbank.
```

SQ Sequence 390 AA;
 Query Match 15.7%; Score 306.5; DB 8; Length 390;
 Best Local Similarity 25.3%; Pred. No. 6e-19;
 Matches 95; Conservative 63; Mismatches 151; Indels 67; Gaps 13;
 QY 24 FAFSVWSDGDTFVRSWDEFRQLKTKLTPVEAGLLRRSDRVLPKLLDAPLLGRVG 83
 DB 26 YMLVKWQDLSEKVVYRRFTIYEFHKTLEMPFIEAGINPENRIIPH-LPAPKWFDDQ 84
 QY 84 RTSRGLARLQLETYSRRLLATAERVARSPITIGFFAPQPLDLEPALPPGSRV-----I 137
 DB 85 RAAEN--RQGLTEYCSLMSLPTKISRCPLDFFKVRPDDUK--LPTDNQTKKPTYL 140
 QY 138 LPTPEQPLSRAAGRLSIHSLAQSLRCLQPFCTQTRDRPFOAQOESLDVLLRHPSGW 197
 DB 141 MPKDGKSTATDITGPIL-----QTYRAINY--EKTSGSEMALSTGDDVVEVVEKSESGW 193
 QY 198 WLVENEDROTANFPAPYLEA-APGQREGGSLGSSGPOFCASRAYESSRADELSVPAG 256
 DB 194 WFCQMKAKR-GWIPASFLPLSDPDETPENY--AGEPYVAIKAYTAVEGDEVSLLEG 250
 QY 257 ARVRVLETSDRGWMLCRYGDRAGLLPAVLLRPEG-----LGA----- 293
 DB 251 EAVEVIHKLLDGNWVIRKDDVTGYFSPMYLQKSGQDVSOAQRIKRGAPPRSSIRNAHS 310
 QY 294 -----LLSGTGFRGGD-----DPAGEARGFP-----EPSQATAPPTVTPRPS 331
 DB 311 IHQSRKRLSODAYRRNSVRFLOQRRRQARPQSPGSPLEERQTRQSKPQPAVPPRPS 370
 QY 332 PGAIQSRCCCTVTRRAL 347
 DB 371 ADLILNRCSESTKRKL 386
 RESULT 13
 ABG20715
 ID ABG20715 standard; protein; 976 AA.
 XX AC ABG20715;
 XX DT 18-FEB-2002 (first entry)
 XX DE Novel human diagnostic protein #20706.
 XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 XX KW food supplement; medical imaging; diagnostic; genetic disorder.
 XX OS Homo sapiens.
 XX PN WO200175067-A2.
 XX PD 11-OCT-2001.
 XX PF 30-MAR-2001; 2001WO-US008631.
 XX PR 31-MAR-2000; 2000US-00540217.
 XX PR 23-AUG-2000; 2000US-00649167.
 XX PA (HYSE-) HYSEQ INC.
 XX PI Drmanac RT, Liu C, Tang YT;
 XX DR WPI; 2001-639362/73.
 XX DR N-PSDB; AAS84902.
 XX New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.
 XX Claim 20; SEQ ID NO 51074; 103pp; English.

CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
 CC sequences. (I) is useful as hybridisation probes, polymerase chain
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
 CC and in recombinant production of (II). The polynucleotides are also used
 CC in diagnostics as expressed sequence tags for identifying expressed
 CC genes. (I) is useful in gene therapy techniques to restore normal
 CC activity of (II) or to treat disease states involving (II). (II) is
 CC useful for generating antibodies against it, detecting or quantitating a
 CC polypeptide in tissue, as molecular weight markers and as a food
 CC supplement. (II) and its binding partners are useful in medical imaging
 CC of sites expressing (II). (I) and (II) are useful for treating disorders
 CC involving aberrant protein expression or biological activity. The
 CC polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
 CC amino acid sequences. The invention. Note: The sequence data for this
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 976 AA;
 Query Match 15.7%; Score 306.5; DB 4; Length 976;
 Best Local Similarity 25.3%; Pred. No. 2e-18;
 Matches 95; Conservative 63; Mismatches 151; Indels 67; Gaps 13;
 QY 24 FAFSVWSDGDTFVRSWDEFRQLKTKLTPVEAGLLRRSDRVLPKLLDAPLLGRVG 83
 DB 612 YMLVKWQDLSEKVVYRRFTIYEFHKTLEMPFIEAGINPENRIIPH-LPAPKWFDDQ 670
 QY 84 RTSRGLARLQLETYSRRLLATAERVARSPITIGFFAPQPLDLEPALPPGSRV-----I 137
 DB 671 RAAEN--RQGLTEYCSLMSLPTKISRCPLDFFKVRPDDUK--LPTDNQTKKPTYL 726
 QY 138 LPTPEQPLSRAAGRLSIHSLAQSLRCLQPFCTQTRDRPFOAQOESLDVLLRHPSGW 197
 DB 727 MPKDGKSTATDITGPIL-----QTYRAINY--EKTSGSEMALSTGDDVVEVVEKSESGW 779
 QY 198 WLVENEDROTANFPAPYLEA-APGQREGGSLGSSGPOFCASRAYESSRADELSVPAG 256
 DB 780 WFCQMKAKR-GWIPASFLPLSDPDETPENY--AGEPYVAIKAYTAVEGDEVSLLEG 836
 QY 257 ARVRVLETSDRGWMLCRYGDRAGLLPAVLLRPEG-----LGA----- 293
 DB 837 EAVEVIHKLLDGNWVIRKDDVTGYFSPMYLQKSGQDVSOAQRIKRGAPPRSSIRNAHS 896
 QY 294 -----LLSGTGFRGGD-----DPAGEARGFP-----EPSQATAPPTVTPRPS 331
 DB 897 IHQSRKRLSODAYRRNSVRFLOQRRRQARPQSPGSPLEERQTRQSKPQPAVPPRPS 956
 QY 332 PGAIQSRCCCTVTRRAL 347
 DB 957 ADLILNRCSESTKRKL 972
 RESULT 14
 ADG36774
 ID ADG36774 standard; protein; 389 AA.
 XX AC ADG36774;
 XX DT 26-FEB-2004 (first entry)
 XX DE Rat P47PHOX protein.
 XX KW rat; autoimmune condition; NADPH oxidase activity;
 KW NADPH oxidase activator; P47PHOX; arthritis; multiple sclerosis; lupus;
 KW autoimmune uveitis; type I diabetes; bronchial asthma; septic arthritis;
 KW cardiovascular disease.
 XX OS Rattus norvegicus.

Db 2 YMFVWKQDLSEKVVRRFTIYEFHKTLEKMPFIEAGAINPENRIIPH-LPAPKWFQDQ 60
QY 81 RVGRTSRGLARQLLETYSRLLATAERVARSPITITGFFAQPQLDLEPALPPGSRV---- 136
Db 61 RAAENHOG-----TLTEYCGTMSLPTKISRCPHLLDFFKVRPDDLK--LPTDNQTKKPE 113
QY 137 --ILPTPEOPLSRAAGRLSIHSLAQSLRCLQPFCTQDTRDRPFOAQOESLDVLLRHP 194
Db 114 TYLMPKDGKSTADITGPILL-----QTYRAIANY--EKTSGSEMALSTGDVVVEVEKSE 166
QY 195 SGHWLVENEDROTAWFPAPVLEEA-APGQREGGPSLGGSGPQFCASRAYESSRADELSV 253
Db 167 SGWFFCQMKAKR-GWIPASLEPLDSPDETEDEPNY--AGEPYVAIKAYTAVEGDEVSL 223
QY 254 PAGARVRVLETSDRGWMLCRYGDRAGLLPAVLLRPEG-----LGA----- 293
Db 224 LEGEAEVVIHKLDDGWWVIRKDDVTGYFFPSMYLQKSGQDVSOAQORQIKRGAPPRSSIRN 283
QY 294 -----LLSGTGFRGSD-----DPAGEARGFP-----EPSQATAPPPTVPT 328
Db 284 VHSIHQSRKRRLSQDAYRRNSVRFLQORRRQARFGPQSGPSPLEEERQTORSKPQPAVPP 343
QY 329 RPSPGAIOQRCCCTVTRAL 347
Db 344 RPSADLILNRCSESTKRKL 362

Search completed: May 28, 2005, 07:40:35
Job time : 165 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 28, 2005, 07:35:33 ; Search time 42 Seconds
(without alignments)
659.400 Million cell updates/sec

Title: US-10-621-113-4

Perfect score: 1950

Sequence: 1 MAGPRYFVSQGAALVQIKR.....RRQRPRGCVDSVPHPPTTEQ 371

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.*

- 1: /cgn2_6/prodata/1/iaa/5A_COMB.pep.*
- 2: /cgn2_6/prodata/1/iaa/5B_COMB.pep.*
- 3: /cgn2_6/prodata/1/iaa/6A_COMB.pep.*
- 4: /cgn2_6/prodata/1/iaa/6B_COMB.pep.*
- 5: /cgn2_6/prodata/1/iaa/PCUTUS_COMB.pep.*
- 6: /cgn2_6/prodata/1/iaa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	307.5	15.8	390	4	US-09-820-005-4
2	307.5	15.8	390	4	US-10-109-856-4
3	278.5	14.3	386	4	US-09-820-005-2
4	278.5	14.3	386	4	US-10-109-856-2
5	153.5	7.9	215	4	US-09-808-701A-21
6	139.5	7.2	560	1	US-07-683-957B-1
7	134	6.9	381	4	US-09-949-016-9555
8	131.5	6.7	614	4	US-09-252-991A-28630
9	130	6.7	532	4	US-09-949-016-7663
10	128.5	6.6	201	4	US-09-949-016-11140
11	128.5	6.6	377	4	US-09-538-092-939
12	128.5	6.6	377	4	US-09-949-016-6693
13	128	6.6	561	4	US-09-252-991A-31113
14	125.5	6.4	1706	4	US-09-252-991A-31760
15	125	6.4	495	4	US-09-252-991A-16992
16	124.5	6.4	462	3	US-08-630-915A-38
17	124.5	6.4	462	4	US-09-879-957-38
18	124.5	6.4	472	4	US-09-252-991A-16723
19	124.5	6.4	520	4	US-09-538-092-1347
20	124	6.4	248	3	US-08-630-915A-40
21	124	6.4	248	4	US-09-879-957-40
22	124	6.4	509	3	US-08-630-915A-194
23	124	6.4	509	4	US-09-879-957-194
24	124	6.4	1676	4	US-09-949-016-7610
25	123.5	6.3	907	4	US-09-252-991A-24114
26	123	6.3	679	4	US-09-252-991A-18857
27	121.5	6.2	917	4	US-09-252-991A-25101

28	119.5	6.1	427	4	US-09-252-991A-30434	Sequence 30434, A
29	119	6.1	554	4	US-09-252-991A-28232	Sequence 28232, A
30	119	6.1	636	4	US-09-252-991A-24902	Sequence 24902, A
31	118.5	6.1	366	4	US-09-252-991A-24338	Sequence 24338, A
32	118.5	6.1	447	3	US-09-199-637A-351	Sequence 351, App
33	118	6.1	847	4	US-09-949-016-6222	Sequence 6222, App
34	118	6.1	885	4	US-09-949-016-7789	Sequence 7789, App
35	118	6.1	1427	4	US-09-252-991A-27005	Sequence 27005, A
36	117.5	6.0	461	4	US-09-252-991A-24717	Sequence 24717, A
37	117	6.0	355	3	US-08-630-915A-192	Sequence 192, App
38	117	6.0	355	4	US-09-879-957-192	Sequence 192, App
39	117	6.0	950	4	US-09-252-991A-25927	Sequence 25927, A
40	116.5	6.0	425	4	US-09-252-991A-17013	Sequence 17013, A
41	115	5.9	526	4	US-09-252-991A-28761	Sequence 28761, A
42	115	5.9	536	4	US-09-252-991A-23495	Sequence 23495, A
43	114.5	5.9	303	4	US-09-538-092-1133	Sequence 1133, App
44	114.5	5.9	486	2	US-08-942-423-2	Sequence 2, Appli
45	114.5	5.9	486	3	US-08-630-915A-26	Sequence 26, Appl

ALIGNMENTS

RESULT 1
US-09-820-005-4
; Sequence 4, Application US/09820005
; Patent No. 6489149
; GENERAL INFORMATION:
; APPLICANT: SHAO, Wei et al
; TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001198
; CURRENT APPLICATION NUMBER: US/09/820,005
; CURRENT FILING DATE: 2001-03-29
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 390
; TYPE: PRT
; ORGANISM: Human
US-09-820-005-4

Query Match	15.8%	Score	307.5	DB	4	Length	390
Best Local Similarity	25.3%	Pred. No.	7.5e-21				
Matches	95	Conservative	63	Mismatches	151	Indels	67
Gaps	13						
QY	24	F A F S V W S D G S D T F V R S W D E F R Q L K K T K E T F P V E A G L L R R S D R V L P K L L D A P L L G R V G	83				
Db	26	Y M F L V K W Q D L S E K V Y R F T E I Y E F H K T L K E M P F I E A G A I N P E N R I I P H - L P A P K W F D G Q	84				
QY	84	R T S R G L A R L Q L L E T Y S R R L L A T A E R V A R S P T T I T G F P A P O P L D L E P A L P P G S R V - - - - - I	137				
Db	85	R A E N - - R G T T E Y C S T L M S L P T K I S R C P H L D F F K V R P D D L K - - L P T D N Q T K K E T Y L	140				
QY	138	L P T P E Q P L S R A A G R L S I H S L E A Q S L R C L Q P F C T Q T D R P F Q A Q A S L D V L L R H P S G W	197				
Db	141	M P K D G K S T A T D I T G P I L - - - - - Q T Y R A I A D Y - - E N T S G S E M A L S T G D V V E V K E S G W	193				
QY	198	W L V E N E D R T A F P P A P Y L E E A - A P Q G R E G G S L G S S G P Q F C A S R A Y E S R A D E L S V P A G	256				
Db	194	W F C Q M K A K R - G M I P A S F L E P L D S P D E T E P E P N Y - - A G E P Y V A I K A Y T A V E G D S V L L E G	250				
QY	257	A R V V L E T S D R G W M L C R Y G D R A G L L P A V L L R P E G - - - - - L G A - - - - - - - - - - - - - - -	293				
Db	251	E A V E V I H K L D G W M V I R K D D V T G Y F P S M Y L Q S G Q D V S Q A Q R I K K A G P P R R S I R N A H S	310				
QY	294	- - - - - L L S G T G F R G G D - - - - - D P A G E A R G F P - - - - - E P S Q A T A P P T V T P T R P S	331				
Db	311	I H Q R S K R L S Q D A Y R N S V R F L Q O R R Q A R P G S P G S P L E E R Q T Q R S K P Q A V P P R P S	370				
QY	332	P G A I O S R C C T V T R R A L 347					


```
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9555
; LENGTH: 381
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-9555

Query Match
Best Local Similarity 24.3%; Score 134; DB 4; Length 381;
Matches 44; Conservative 20; Mismatches 57; Indels 60; Gaps 6;

QY 179 FOAQESLDV-----LLRHPSGWLVENEDROTAMFPAPYLE-----216
Db 14 YTAQDQELDIKKNERLMLLDDSKTWVRVNAANRTGVPSNYVERKNSLKKGLVKNLK 73
QY 217 -----EAPGQREGG-PSLGS-----PQCASRAYESSRADELS 252
Db 74 DTLGLGKTRKTSADSTPTDAEYPANGSGADRIYDLNIPAF-VKPAYVAEREDELS 132
QY 253 VPAGARVRVLETSDRGWMLCRYGDRAGLLPAVLLRPEGLGALLSGTGRGGDDPAGEARG 312
Db 133 LVKGSRTVMEKCSGWRGSGNQGIFWFSNYVLEE-----VDEAAAESPS 179
QY 313 F 313
Db 180 F 180

RESULT 8
US-09-252-991A-28630
; Sequence 28630, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 28630
; LENGTH: 614
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-28630

Query Match
Best Local Similarity 24.2%; Score 131.5; DB 4; Length 614;
Matches 80; Conservative 21; Mismatches 109; Indels 121; Gaps 17;

QY 126 LEPALP---PGSRVILPTPEOPLSRAAGRLSIHSLEAQSRLCLOPFCTQD-----TRDRP 178
Db 88 VDPAPADPDCGRLRRDRPRRPPANRRHRRTRHVPG-----APPVPLADGAAERRP 142
QY 179 FQ-----AQAESLDVLLRHPSGWLVENEDROTAMFPAPYLEEAPGOGREGGSLGSSG 234
Db 143 GQSPWRNRQGGADRII-HPPG-----RPARLRACLPAVVVRHGPARGDRPWPQG-GQPA 195
QY 235 PQFCASRAYESSRADELSVPAGARVRVLETSDRGWMLCRYGD-----RAG 279

; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9555
; LENGTH: 381
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-9555

Query Match
Best Local Similarity 24.3%; Score 134; DB 4; Length 381;
Matches 44; Conservative 20; Mismatches 57; Indels 60; Gaps 6;

QY 179 FOAQESLDV-----LLRHPSGWLVENEDROTAMFPAPYLE-----216
Db 14 YTAQDQELDIKKNERLMLLDDSKTWVRVNAANRTGVPSNYVERKNSLKKGLVKNLK 73
QY 217 -----EAPGQREGG-PSLGS-----PQCASRAYESSRADELS 252
Db 74 DTLGLGKTRKTSADSTPTDAEYPANGSGADRIYDLNIPAF-VKPAYVAEREDELS 132
QY 253 VPAGARVRVLETSDRGWMLCRYGDRAGLLPAVLLRPEGLGALLSGTGRGGDDPAGEARG 312
Db 133 LVKGSRTVMEKCSGWRGSGNQGIFWFSNYVLEE-----VDEAAAESPS 179
QY 313 F 313
Db 180 F 180

RESULT 8
US-09-252-991A-28630
; Sequence 28630, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 28630
; LENGTH: 614
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-28630

Query Match
Best Local Similarity 24.2%; Score 131.5; DB 4; Length 614;
Matches 80; Conservative 21; Mismatches 109; Indels 121; Gaps 17;

QY 126 LEPALP---PGSRVILPTPEOPLSRAAGRLSIHSLEAQSRLCLOPFCTQD-----TRDRP 178
Db 88 VDPAPADPDCGRLRRDRPRRPPANRRHRRTRHVPG-----APPVPLADGAAERRP 142
QY 179 FQ-----AQAESLDVLLRHPSGWLVENEDROTAMFPAPYLEEAPGOGREGGSLGSSG 234
Db 143 GQSPWRNRQGGADRII-HPPG-----RPARLRACLPAVVVRHGPARGDRPWPQG-GQPA 195
QY 235 PQFCASRAYESSRADELSVPAGARVRVLETSDRGWMLCRYGD-----RAG 279

; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7663
; LENGTH: 532
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-7663

Query Match
Best Local Similarity 26.6%; Score 130; DB 4; Length 532;
Matches 83; Conservative 21; Mismatches 130; Indels 78; Gaps 14;

QY 89 LARIQLLETYSRRLLATAERVARSPTI--TGFFAPQPLDLEPALP--PGSRVIL---PTP 141
Db 6 LTSLALVLSALGALLGTEALRAEPAVGTSGLIIFREDLDWPPGIPQEPCLCLVALGDSNG 65
QY 142 EQPLSRAAGRLSIHSLEAQSIL-----RCLOPCTQDTRDRPFOAQAESLDVLL 191
Db 66 SSSPL-RVVGALSAY--EQAFLGAVQARAWGFRDLATFGVNCNTGDRQALPSRLRLGAWL 122
QY 192 RHPSGWLVENEDROTAMFPAPYLEEAPGOGREGGSLG-----SSGPOFCASRAYES 245
Db 123 RDPGQRLVVLHLEEVWTEPTPSLRFQEPFGGAGPPELALLVYPGGPEVTVTRA--- 179
QY 246 SRADELSVPAGARVRVLETSDRGWMLCRYGDRAGLLPAVLLRPEGLGALLSGTGRGGDD 305
Db 180 -----GLP-GAQS-----LCPSRDRVLVLAV-----DR 202
QY 306 PAGEARGPEPSQATAPPTVTPRPSGAIQS-----RCCTVTRRALLERRRPRQGRP- 357
Db 203 PAGAWRG--SGIALTLQPRGEDSRLSTARLQALLFGDDHRCFTRMTPALLLPRSEAPL 260
QY 358 --RGCVDSPVHP 367
Db 261 PAHQDLTVPEP 272

RESULT 10
US-09-949-016-11140
; Sequence 11140, Application US/09949016
```

```

; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11140
; LENGTH: 201
; TYPE: PRT
; ORGANISM: Human
; US-09-949-016-11140

Query Match      6.6%; Score 128.5; DB 4; Length 201;
Best Local Similarity 24.6%; Pred. No. 0.00032;
Matches 44; Conservative 20; Mismatches 52; Indels 63; Gaps 6;

QY 179 FOAQOESLDV-----LLRHPGWLWVENEDRQTAWFPAPYLE-----E 217
Db 14 YVAQOQELDIKKNERLWLLDDSKSWVRNSNMKTGFVPSNYVERKNSARKASIVKNLK 73
QY 218 AAPGQGR-EGGPSLGSSGPQFCASRA-----YESSRADELSV 253
Db 74 DTLGIGKVKRKPSVDPD-----ASPADDSPVDPGERLYDLNMPAYVKFNMAEREDELSL 128
QY 254 PAGARVRVLETSDRGWLWLCRYGDRAGLLPAVLLRPEGLGALLSGTGFRGDDPAGEARG 312
Db 129 IKGTKVIVMEKCDGWMRGSYNGQGVWFPNSYVTEE-----GDSPLGDHVG 174

RESULT 11
US-09-538-092-939
; Sequence 939, Application US/09538092
; Patent No. 6753314
; GENERAL INFORMATION:
; APPLICANT: Mansfield, Traci A.
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
; FILE REFERENCE: 15966-542
; CURRENT APPLICATION NUMBER: US/09/538,092
; CURRENT FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/178,965
; PRIOR FILING DATE: 2000-02-01
; NUMBER OF SEQ ID NOS: 1387
; SOFTWARE: CuraPatSeqFormatter Version 0.9
; SEQ ID NO 939
; LENGTH: 377
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (0)..(0)
; OTHER INFORMATION: Polypeptide Accession Number P16333
US-09-538-092-939

Query Match      6.6%; Score 128.5; DB 4; Length 377;
Best Local Similarity 24.6%; Pred. No. 0.00077;
Matches 44; Conservative 20; Mismatches 52; Indels 63; Gaps 6;

QY 179 FOAQOESLDV-----LLRHPGWLWVENEDRQTAWFPAPYLE-----E 217
Db 13 YVAQOQELDIKKNERLWLLDDSKSWVRNSNMKTGFVPSNYVERKNSARKASIVKNLK 72

; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11140
; LENGTH: 201
; TYPE: PRT
; ORGANISM: Human
; US-09-949-016-11140

Query Match      6.6%; Score 128.5; DB 4; Length 377;
Best Local Similarity 24.6%; Pred. No. 0.00077;
Matches 44; Conservative 20; Mismatches 52; Indels 63; Gaps 6;

QY 179 FOAQOESLDV-----LLRHPGWLWVENEDRQTAWFPAPYLE-----E 217
Db 13 YVAQOQELDIKKNERLWLLDDSKSWVRNSNMKTGFVPSNYVERKNSARKASIVKNLK 72
QY 218 AAPGQGR-EGGPSLGSSGPQFCASRA-----YESSRADELSV 253
Db 73 DTLGIGKVKRKPSVDPD-----ASPADDSPVDPGERLYDLNMPAYVKFNMAEREDELSL 127
QY 254 PAGARVRVLETSDRGWLWLCRYGDRAGLLPAVLLRPEGLGALLSGTGFRGDDPAGEARG 312
Db 128 IKGTKVIVMEKCDGWMRGSYNGQGVWFPNSYVTEE-----GDSPLGDHVG 173

RESULT 12
US-09-949-016-6693
; Sequence 6693, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6693
; LENGTH: 377
; TYPE: PRT
; ORGANISM: Human
; US-09-949-016-6693

Query Match      6.6%; Score 128.5; DB 4; Length 377;
Best Local Similarity 24.6%; Pred. No. 0.00077;
Matches 44; Conservative 20; Mismatches 52; Indels 63; Gaps 6;

QY 179 FOAQOESLDV-----LLRHPGWLWVENEDRQTAWFPAPYLE-----E 217
Db 13 YVAQOQELDIKKNERLWLLDDSKSWVRNSNMKTGFVPSNYVERKNSARKASIVKNLK 72
QY 218 AAPGQGR-EGGPSLGSSGPQFCASRA-----YESSRADELSV 253
Db 73 DTLGIGKVKRKPSVDPD-----ASPADDSPVDPGERLYDLNMPAYVKFNMAEREDELSL 127
QY 254 PAGARVRVLETSDRGWLWLCRYGDRAGLLPAVLLRPEGLGALLSGTGFRGDDPAGEARG 312
Db 128 IKGTKVIVMEKCDGWMRGSYNGQGVWFPNSYVTEE-----GDSPLGDHVG 173

RESULT 13
US-09-252-991A-31113
; Sequence 3113, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 31113
; LENGTH: 561
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
; US-09-252-991A-31113

```


Search completed: May 28, 2005, 07:45:02
Job time : 43 secs



This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 28, 2005, 07:41:29 ; Search time 141 Seconds
(without alignments)
907.693 Million cell updates/sec

Title: US-10-621-113-4

Perfect score: 1950

Sequence: 1 MAGPRYPVSVQGAALVQIKR.....RRQGRPGCVDSVPHPTTEQ 371

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1462099 seqs, 344972447 residues

Total number of hits satisfying chosen parameters: 1462099

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
- 19: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 20: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1950	100.0	371	15 US-10-621-113-4	Sequence 4, Appli
2	1937.5	99.4	376	15 US-10-621-113-6	Sequence 6, Appli
3	1934.5	99.2	370	15 US-10-621-113-2	Sequence 2, Appli
4	1922	98.6	375	15 US-10-621-113-8	Sequence 8, Appli
5	309.5	15.9	390	15 US-10-437-427-7	Sequence 7, Appli
6	307.5	15.8	390	14 US-10-109-856-4	Sequence 4, Appli
7	307.5	15.8	390	15 US-10-418-036-18	Sequence 18, Appli
8	307.5	15.8	390	15 US-10-437-427-6	Sequence 6, Appli
9	307.5	15.8	390	16 US-10-767-341-4	Sequence 4, Appli
10	307	15.7	389	15 US-10-437-427-4	Sequence 4, Appli
11	306.5	15.7	390	16 US-10-755-889-451	Sequence 451, App
12	306	15.7	389	15 US-10-437-427-2	Sequence 2, Appli
13	278.5	14.3	386	14 US-10-109-856-2	Sequence 2, Appli

14	278.5	14.3	386	16	US-10-767-341-2	Sequence 2, Appli
15	245.5	12.6	1138	15	US-10-161-927-60	Sequence 60, Appli
16	211	10.8	1054	15	US-10-094-749-2576	Sequence 2576, Ap
17	163	8.4	268	9	US-09-764-868-721	Sequence 721, App
18	158	8.1	204	11	US-09-764-875-874	Sequence 874, App
19	153.5	7.9	215	9	US-09-808-701A-21	Sequence 21, Appli
20	153.5	7.9	215	14	US-10-233-131-21	Sequence 21, Appli
21	153.5	7.9	215	15	US-10-240-145-73	Sequence 73, Appli
22	146	7.5	968	15	US-10-291-265-281	Sequence 281, App
23	143	7.3	134	14	US-10-202-724-4	Sequence 4, Appli
24	141.5	7.3	769	15	US-10-369-493-4212	Sequence 4212, Ap
25	139.5	7.2	560	14	US-10-162-223-7	Sequence 7, Appli
26	139.5	7.2	560	15	US-10-366-345-53	Sequence 53, Appli
27	139.5	7.2	561	9	US-09-813-398-23	Sequence 23, Appli
28	139	7.1	370	10	US-09-946-374-315	Sequence 315, App
29	139	7.1	370	13	US-10-052-586-350	Sequence 350, App
30	139	7.1	370	14	US-10-174-590-350	Sequence 350, App
31	139	7.1	370	14	US-10-176-758-350	Sequence 350, App
32	139	7.1	370	14	US-10-175-737-350	Sequence 350, App
33	139	7.1	370	14	US-10-174-581-350	Sequence 350, App
34	139	7.1	370	14	US-10-176-483-350	Sequence 350, App
35	139	7.1	370	14	US-10-176-749-350	Sequence 350, App
36	139	7.1	370	14	US-10-176-914-350	Sequence 350, App
37	139	7.1	370	14	US-10-173-706-350	Sequence 350, App
38	139	7.1	370	14	US-10-175-738-350	Sequence 350, App
39	139	7.1	370	14	US-10-175-752-350	Sequence 350, App
40	139	7.1	370	14	US-10-176-482-350	Sequence 350, App
41	139	7.1	370	14	US-10-176-757-350	Sequence 350, App
42	139	7.1	370	14	US-10-176-913-350	Sequence 350, App
43	139	7.1	370	14	US-10-180-552-350	Sequence 350, App
44	139	7.1	370	14	US-10-180-557-350	Sequence 350, App
45	139	7.1	370	14	US-10-180-557-350	Sequence 350, App

ALIGNMENTS

RESULT 1
US-10-621-113-4
; Sequence 4, Application US/10621113
; Publication No. US20040091466A1
; GENERAL INFORMATION:
; APPLICANT: Lambeth, J. David
; APPLICANT: Cheng, Guangjie
; TITLE OF INVENTION: Regulatory Protein For Nox Enzymes
; FILE REFERENCE: 05501-0202 (43150-287577)
; CURRENT APPLICATION NUMBER: US/10/621,113
; CURRENT FILING DATE: 2003-07-16
; PRIOR APPLICATION NUMBER: US 60/405,647
; PRIOR FILING DATE: 2002-08-23
; PRIOR APPLICATION NUMBER: US 60/396,170
; PRIOR FILING DATE: 2002-07-16
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: Patent version 3.1
; SEQ ID NO 4
; LENGTH: 371
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-621-113-4

Query Match	100.0%	Score	1950;	DB	15;	Length	371;
Best Local Similarity	100.0%	Pred. No.	8.1e-148;				
Matches	371;	Conservative	0;	Mismatches	0;	Indels	0;
Gaps	0;						
QY	1	MAGPRYPVSVQGAALVQIKRLQTFAPSVWSDGSDTFVRSWDEFPRLKKTLETPEVEA	60				
Db	1	MAGPRYPVSVQGAALVQIKRLQTFAPSVWSDGSDTFVRSWDEFPRLKKTLETPEVEA	60				
QY	61	GLLRSDRVLPLKLLDAPLGRVGRSGLARLQLLETYSRRLLATAEVARSPITIGFFA	120				
Db	61	GLLRSDRVLPLKLLDAPLGRVGRSGLARLQLLETYSRRLLATAEVARSPITIGFFA	120				
QY	121	POPLDLEPALPGSRVILPTPEEQPLSRAAGRLSIHSLEAQLRCLQPFCTQTRDRPFQ	180				

Db 121 PQLDLEPALPGSRVILPTPEEQPLSRAAGRLSIHSLEAQSLRCLQPCTQDTRDRPFQ 180
QY 181 AQAQESLDVLRHPSGWLWVENEDRQTAWFPAPYLEEAPGQREGGPSLGSSGPFQFCAS 240
Db 181 AQAQESLDVLRHPSGWLWVENEDRQTAWFPAPYLEEAPGQREGGPSLGSSGPFQFCAS 240
QY 241 RAYESSRADELSPAGARVRVLETSDRGWLWCYRQDRAGLLPAVLLRPEGLGALLSGTGF 300
Db 241 RAYESSRADELSPAGARVRVLETSDRGWLWCYRQDRAGLLPAVLLRPEGLGALLSGTGF 300
QY 301 RCGDDPAGEARGFPPEPSQATAPPTVTPRPSGAIQSRCTVTRALERRRRQGRPRGC 360
Db 301 RCGDDPAGEARGFPPEPSQATAPPTVTPRPSGAIQSRCTVTRALERRRRQGRPRGC 360
QY 361 VDSVPHPPTTEQ 371
Db 361 VDSVPHPPTTEQ 371

RESULT 2

US-10-621-113-6

; Sequence 6, Application US/10621113
; Publication No. US20040091466A1

; GENERAL INFORMATION:

; APPLICANT: Lambeth, J. David

; TITLE OF INVENTION: Regulatory Protein For Nox Enzymes

; FILE REFERENCE: 05501-0202 (43150-287577)

; CURRENT APPLICATION NUMBER: US/10/621,113

; CURRENT FILING DATE: 2003-07-16

; PRIOR APPLICATION NUMBER: US 60/405,647

; PRIOR FILING DATE: 2002-08-23

; PRIOR APPLICATION NUMBER: US 60/396,170

; PRIOR FILING DATE: 2002-07-16

; NUMBER OF SEQ ID NOS: 11

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 6

; LENGTH: 376

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-621-113-6

Query Match 99.4%; Score 1937.5; DB 15; Length 376;

Best Local Similarity 98.7%; Pred. No. 8.3e-147; Mismatches 0; Indels 5; Gaps 1;

Matches 371; Conservative 0; Mismatches 0; Indels 5; Gaps 1;

QY 1 MAGPRYPVSQGAALVQIKRLQTFAFSVRWSDGSDTFVRRSWDEFRQLKTKLTFPVEA 60
Db 1 MAGPRYPVSQGAALVQIKRLQTFAFSVRWSDGSDTFVRRSWDEFRQLKTKLTFPVEA 60
QY 61 GLLRRSDRVLPKL-----LDAPLLGRVGTSGRLARLQLLETYSRRLATAERVARSPTI 115
Db 61 GLLRRSDRVLPKLQQAISLDAPLLGRVGTSGRLARLQLLETYSRRLATAERVARSPTI 120
QY 116 TGFFAPQPLDLEPALPGSRVILPTPEEQPLSRAAGRLSIHSLEAQSLRCLQPCTQDTR 175
Db 121 TGFFAPQPLDLEPALPGSRVILPTPEEQPLSRAAGRLSIHSLEAQSLRCLQPCTQDTR 180
QY 176 DRPFOAQESLDVLRHPSGWLWVENEDRQTAWFPAPYLEEAPGQREGGPSLGSSGP 235
Db 181 DRPFOAQESLDVLRHPSGWLWVENEDRQTAWFPAPYLEEAPGQREGGPSLGSSGP 240
QY 236 QFCASRAYESSRADELSPAGARVRVLETSDRGWLWCYRQDRAGLLPAVLLRPEGLGALL 295
Db 241 QFCASRAYESSRADELSPAGARVRVLETSDRGWLWCYRQDRAGLLPAVLLRPEGLGALL 300
QY 296 SCTGRGDDPAGEARGFPPEPSQATAPPTVTPRPSGAIQSRCTVTRALERRRRQGR 355
Db 301 SCTGRGDDPAGEARGFPPEPSQATAPPTVTPRPSGAIQSRCTVTRALERRRRQGR 360
QY 356 RPRGCVDSVPHPPTTEQ 371
Db 356 RPRGCVDSVPHPPTTEQ 371

Db 361 RPRGCVDSVPHPPTTEQ 376

RESULT 3

US-10-621-113-2

; Sequence 2, Application US/10621113

; Publication No. US20040091466A1

; GENERAL INFORMATION:

; APPLICANT: Lambeth, J. David

; APPLICANT: Cheng, Guangjie

; TITLE OF INVENTION: Regulatory Protein For Nox Enzymes

; FILE REFERENCE: 05501-0202 (43150-287577)

; CURRENT APPLICATION NUMBER: US/10/621,113

; CURRENT FILING DATE: 2003-07-16

; PRIOR APPLICATION NUMBER: US 60/405,647

; PRIOR FILING DATE: 2002-08-23

; PRIOR APPLICATION NUMBER: US 60/396,170

; PRIOR FILING DATE: 2002-07-16

; NUMBER OF SEQ ID NOS: 11

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 2

; LENGTH: 370

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-621-113-2

Query Match 99.2%; Score 1934.5; DB 15; Length 370;

Best Local Similarity 99.7%; Pred. No. 1.4e-146; Mismatches 0; Indels 1; Gaps 1;

Matches 370; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 MAGPRYPVSQGAALVQIKRLQTFAFSVRWSDGSDTFVRRSWDEFRQLKTKLTFPVEA 60
Db 1 MAGPRYPVSQGAALVQIKRLQTFAFSVRWSDGSDTFVRRSWDEFRQLKTKLTFPVEA 59
QY 61 GLLRRSDRVLPKLDAPLLGRVGTSGRLARLQLLETYSRRLATAERVARSPITGFFA 120
Db 60 GLLRRSDRVLPKLDAPLLGRVGTSGRLARLQLLETYSRRLATAERVARSPITGFFA 119
QY 121 PQLDLEPALPGSRVILPTPEEQPLSRAAGRLSIHSLEAQSLRCLQPCTQDTRDRPFQ 180
Db 120 PQLDLEPALPGSRVILPTPEEQPLSRAAGRLSIHSLEAQSLRCLQPCTQDTRDRPFQ 179
QY 181 AQAQESLDVLRHPSGWLWVENEDRQTAWFPAPYLEEAPGQREGGPSLGSSGPFQFCAS 240
Db 180 AQAQESLDVLRHPSGWLWVENEDRQTAWFPAPYLEEAPGQREGGPSLGSSGPFQFCAS 239
QY 241 RAYESSRADELSPAGARVRVLETSDRGWLWCYRQDRAGLLPAVLLRPEGLGALLSGTGF 300
Db 240 RAYESSRADELSPAGARVRVLETSDRGWLWCYRQDRAGLLPAVLLRPEGLGALLSGTGF 299
QY 301 RCGDDPAGEARGFPPEPSQATAPPTVTPRPSGAIQSRCTVTRALERRRRQGRPRGC 360
Db 300 RCGDDPAGEARGFPPEPSQATAPPTVTPRPSGAIQSRCTVTRALERRRRQGRPRGC 359
QY 361 VDSVPHPPTTEQ 371
Db 360 VDSVPHPPTTEQ 370

RESULT 4
US-10-621-113-8
; Sequence 8, Application US/10621113
; Publication No. US20040091466A1

; GENERAL INFORMATION:
; APPLICANT: Lambeth, J. David
; APPLICANT: Cheng, Guangjie
; TITLE OF INVENTION: Regulatory Protein For Nox Enzymes
; FILE REFERENCE: 05501-0202 (43150-287577)
; CURRENT APPLICATION NUMBER: US/10/621,113
; CURRENT FILING DATE: 2003-07-16
; PRIOR APPLICATION NUMBER: US 60/405,647
; PRIOR FILING DATE: 2002-08-23
; PRIOR APPLICATION NUMBER: US 60/396,170

		PRIOR FILING DATE: 2002-07-16			
		NUMBER OF SEQ ID NOS: 11			
		SOFTWARE: PatentIn version 3.1			
		SEQ ID NO 8			
		LENGTH: 375			
		TYPE: PRT			
		ORGANISM: Homo sapiens			
		US-10-621-113-8			
		Query Match 98.6%; Score 1922; DB 15; Length 375;			
		Best Local Similarity 98.4%; Pred. No. 1.4e-145;			
		Matches 370; Conservative 0; Mismatches 0; Indels 6; Gaps 2;			
Qy	1	MAGPRYPVSVOGAAVQIKLQTFAPSVRWSDGSDTFVRSWDFRQLKTKLKEFPVEA	60		
Db	1	MAGPRYPVSVOGAAVQIKLQTFAPSVRWSDGSDTFVRSWDFRQL-NTLKETPVEA	59		
Qy	61	GLLRSDRVLPKL-----LDAPILGRVGRTRGLARLQILETYSRRLLATAAERVSPTI	115		
Db	60	GLLRSDRVLPKLQGAQLDAPILGRVGRTRGLARLQILETYSRRLLATAAERVSPTI	119		
Qy	116	TGFFAPQDLEPALPGSVILPTPEEQPLSRAAGRLSIHSLEAQSLRCLQPFCTQDTR	175		
Db	120	TGFFAPQDLEPALPGSVILPTPEEQPLSRAAGRLSIHSLEAQSLRCLQPFCTQDTR	179		
Qy	176	DRPFOAQESLDVLLRHPSGWLVENEDRQTAFWPAPYLEEAAPGQREGPSLSSGP	235		
Db	180	DRPFOAQESLDVLLRHPSGWLVENEDRQTAFWPAPYLEEAAPGQREGPSLSSGP	239		
Qy	236	QFCASRAYESSRADELSVPAGARVRLTSDRGWMLCRYGDRAGLLPAVLLRPEGLGALL	295		
Db	240	QFCASRAYESSRADELSVPAGARVRLTSDRGWMLCRYGDRAGLLPAVLLRPEGLGALL	299		
Qy	296	SGTGFGRGDDPAGARGFPSPQATAPPVTPRPSGAIQSRCCCTVTREALRRPRQG	355		
Db	300	SGTGFGRGDDPAGARGFPSPQATAPPVTPRPSGAIQSRCCCTVTREALRRPRQG	359		
Qy	356	RPRGCVDSVPHPPTTEQ	371		
Db	360	RPRGCVDSVPHPPTTEQ	375		
		RESULT 5			
		US-10-437-427-7			
		Sequence 7, Application US/10437427			
		Publication No. US200400901A1			
		GENERAL INFORMATION:			
		APPLICANT: Rikard Holmdahl			
		TITLE OF INVENTION: Autoimmune Conditions and NADPH Oxidase			
		FILE REFERENCE: 11145-024001			
		CURRENT APPLICATION NUMBER: US/10/437,427			
		PRIOR FILING DATE: 2003-05-13			
		PRIOR APPLICATION NUMBER: US 60/380,904			
		PRIOR FILING DATE: 2002-05-13			
		PRIOR APPLICATION NUMBER: US 60/429,609			
		NUMBER OF SEQ ID NOS: 8			
		SOFTWARE: FastSeq for Windows Version 4.0			
		SEQ ID NO 7			
		LENGTH: 390			
		TYPE: PRT			
		ORGANISM: Homo sapiens			
		US-10-437-427-7			
		Query Match 15.9%; Score 309.5; DB 15; Length 390;			
		Best Local Similarity 25.5%; Pred. No. 2.1e-16;			
		Matches 96; Conservative 62; Mismatches 151; Indels 67; Gaps 13;			
Qy	24	FAPSVRWSDGSDTFVRSWDFRQLKTKLKEFPVAGLLRRSDRVLPKLLDAPLLGRVG	83		
Db	26	YMFVKWQDLSEKVVYRRETFEYFHKLEMPFIAGAINPENRIIPH-LPAPKWFQDQ	84		
		us-10-621-113-4			
		PRIOR FILING DATE: 2002-07-16			
		NUMBER OF SEQ ID NOS: 11			
		SOFTWARE: PatentIn version 3.1			
		SEQ ID NO 8			
		LENGTH: 375			
		TYPE: PRT			
		ORGANISM: Homo sapiens			
		US-10-621-113-8			
		Query Match 98.6%; Score 1922; DB 15; Length 375;			
		Best Local Similarity 98.4%; Pred. No. 1.4e-145;			
		Matches 370; Conservative 0; Mismatches 0; Indels 6; Gaps 2;			
Qy	1	MAGPRYPVSVOGAAVQIKLQTFAPSVRWSDGSDTFVRSWDFRQLKTKLKEFPVEA	60		
Db	1	MAGPRYPVSVOGAAVQIKLQTFAPSVRWSDGSDTFVRSWDFRQL-NTLKETPVEA	59		
Qy	61	GLLRSDRVLPKL-----LDAPILGRVGRTRGLARLQILETYSRRLLATAAERVSPTI	115		
Db	60	GLLRSDRVLPKLQGAQLDAPILGRVGRTRGLARLQILETYSRRLLATAAERVSPTI	119		
Qy	116	TGFFAPQDLEPALPGSVILPTPEEQPLSRAAGRLSIHSLEAQSLRCLQPFCTQDTR	175		
Db	120	TGFFAPQDLEPALPGSVILPTPEEQPLSRAAGRLSIHSLEAQSLRCLQPFCTQDTR	179		
Qy	176	DRPFOAQESLDVLLRHPSGWLVENEDRQTAFWPAPYLEEAAPGQREGPSLSSGP	235		
Db	180	DRPFOAQESLDVLLRHPSGWLVENEDRQTAFWPAPYLEEAAPGQREGPSLSSGP	239		
Qy	236	QFCASRAYESSRADELSVPAGARVRLTSDRGWMLCRYGDRAGLLPAVLLRPEGLGALL	295		
Db	240	QFCASRAYESSRADELSVPAGARVRLTSDRGWMLCRYGDRAGLLPAVLLRPEGLGALL	299		
Qy	296	SGTGFGRGDDPAGARGFPSPQATAPPVTPRPSGAIQSRCCCTVTREALRRPRQG	355		
Db	300	SGTGFGRGDDPAGARGFPSPQATAPPVTPRPSGAIQSRCCCTVTREALRRPRQG	359		
Qy	356	RPRGCVDSVPHPPTTEQ	371		
Db	360	RPRGCVDSVPHPPTTEQ	375		
		RESULT 6			
		US-10-109-856-4			
		Sequence 4, Application US/10109856			
		Publication No. US20030166185A1			
		GENERAL INFORMATION:			
		APPLICANT: SHAO, Wei et al.			
		TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC			
		FILE REFERENCE: CLO01198DIV			
		CURRENT APPLICATION NUMBER: US/10/109,856			
		PRIOR FILING DATE: 2002-04-01			
		PRIOR APPLICATION NUMBER: 09/820,005			
		PRIOR FILING DATE: 2001-03-29			
		NUMBER OF SEQ ID NOS: 4			
		SOFTWARE: FastSeq for Windows Version 4.0			
		SEQ ID NO 4			
		LENGTH: 390			
		TYPE: PRT			
		ORGANISM: Homo sapien			
		US-10-109-856-4			
		Query Match 15.8%; Score 307.5; DB 14; Length 390;			
		Best Local Similarity 25.3%; Pred. No. 3.1e-16;			
		Matches 95; Conservative 63; Mismatches 151; Indels 67; Gaps 13;			
Qy	24	FAPSVRWSDGSDTFVRSWDFRQLKTKLKEFPVAGLLRRSDRVLPKLLDAPLLGRVG	83		
Db	26	YMFVKWQDLSEKVVYRRETFEYFHKLEMPFIAGAINPENRIIPH-LPAPKWFQDQ	84		
Qy	84	RTSRGLARLQILETYSRRLLATAAERVSPTITGFPAPQDLEPALPGSRV-----I	137		
Db	85	RAEN--RGTLTEYCVSTLMSLPTKISRCPHLDDFKVPRDDLK--LPTDNQTKKPETYL	140		
Qy	138	LPTPEEQPLSRAAGRLSIHSLEAQSLRCLQPFCTQDTRDRPFOAQAESLDVLLRHPSGW	197		
Db	141	MPKDGKSTATDITGPIL-----QTVRAIADY--EKTSGSEMAISGDDVVVEKSES	193		
Qy	198	WLVENEDRQTAFWPAPYLEEA-APQGREGPSLSSGGPFCASRAYESSRADELSVPAG	256		
Db	194	WFCOMKAKR-GWIPASFLPLSDPDETPENY--AGEPYVAIKAYTAVEGDEVSLLEG	250		
Qy	257	ARVRLTSDRGWMLCRYGDRAGLLPAVLLRPEG-----LGA-----	293		
Db	251	EAVEVIHKLDDGWWIRKDDVTGYFSPMYLQSGQDVSAQQRQIKRGAPPRSSIRNVHS	310		
Qy	294	-----LLSGTGFRGGD-----DPAGEARGFP-----EPSQATAPPTVTPRPS	331		
Db	311	IHQSRKRLSQDAYRNSVRFLQORRRQARPGSPGSPLEERQOTQRSKPEQPAVPPRPS	370		
Qy	332	PGAIQSRCCCTVTTRAL	347		
Db	371	ADILNRCSSESTKRKL	386		

```
Db 311 IHQSRKRLSDAYRNSVRFLOQRROARPGSPGSPLEBERQTRSKQPAPVPPRS 370
QY 332 PGAIQSRCTVTTRAL 347
Db 371 ADLILNRCSESTKRKL 386
```

RESULT 7

```
US-10-418-036-18
; Sequence 18, Application US/10418036
; Publication No. US20030225117A1
; GENERAL INFORMATION:
; APPLICANT: Gronberg, Alvar
; APPLICANT: Wikstrom, Per
; TITLE OF INVENTION: NEW USE
; FILE REFERENCE: 13425-110001
; CURRENT APPLICATION NUMBER: US/10/418,036
; CURRENT FILING DATE: 2003-04-17
; PRIOR APPLICATION NUMBER: SE 0201152-6
; PRIOR FILING DATE: 2002-04-17
; PRIOR APPLICATION NUMBER: US 60/410,626
; PRIOR FILING DATE: 2002-09-13
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 390
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-418-036-18
```

```
Query Match 15.8%; Score 307.5; DB 15; Length 390;
Best Local Similarity 25.3%; Pred. No. 3.1e-16;
Matches 95; Conservative 63; Mismatches 151; Indels 67; Gaps 13;
```

```
QY 24 FAFSVRWSDGSDTFVRRSWDEFRQLKTLKETFPVEAGLLRRSDRVLPKLLDAPLLGRVG 83
Db 26 YMFVKWQDLSEKVVYRRFTIYEFHKLTKEMFPIEAGAINPENRIIPH-LPAPKWFQDG 84
QY 84 RTSRGLARQLLETYSRRLATAERVARSPITITGFFAPQPLDLEPALPPGSRV-----I 137
Db 85 RAAEN--RQGLTEYCVSTLMSLPTKISCPHLLDFKVRPDDLK--LPTDNQTKKPEYTL 140
QY 138 LPTPEEQLSRAAGRLSIHSLAQSLRCLQPCQTODTRDRPFQAQAESLDVLLRHPGSM 197
Db 141 MPKDGKSTATDITGPILL-----QTYRAIADY--EKTSGSEWALSTGDVVEVKESEGW 193
QY 198 WLVENEDQTAFFAPYLEEA-APGQREGGPGSLGSSGPFQFCASRAYESSRADELSPVAG 256
Db 194 WFCQMKAKR-GWIPASFLEPLDSPDETDPENY--AGEPYVAIKAYTAVBGDEVSLLEG 250
QY 257 ARVRVLETSDRGWLCRYGDRAGLLPAVLLRPEG-----LGA----- 293
Db 251 EAVEVIHKLDDGWVIRKDDVTGYFFPSMYLQSGQDVSAQORQIKRGAPRRSSIRNAHS 310
QY 294 -----LLSGTGFRRGD-----DPAGEARGFP-----EFSQATAPPTVTPRPS 331
Db 251 EAVEVIHKLDDGWVIRKDDVTGYFFPSMYLQSGQDVSAQORQIKRGAPRRSSIRNAHS 310
QY 311 IHQSRKRLSDAYRNSVRFLOQRROARPGSPGSPLEBERQTRSKQPAPVPPRS 370
QY 332 PGAIQSRCTVTTRAL 347
Db 371 ADLILNRCSESTKRKL 386
```

RESULT 8

```
US-10-437-427-6
; Sequence 6, Application US/10437427
; Publication No. US2004000901A1
; GENERAL INFORMATION:
; APPLICANT: Rikard Holmdahl
; APPLICANT: Peter Olofsson
; TITLE OF INVENTION: Autoimmune Conditions and NADPH Oxidase
; TITLE OF INVENTION: Defects
```

```
; FILE REFERENCE: 11145-024001
; CURRENT APPLICATION NUMBER: US/10/437,427
; CURRENT FILING DATE: 2003-05-13
; PRIOR APPLICATION NUMBER: US 60/380,904
; PRIOR FILING DATE: 2002-05-13
; PRIOR APPLICATION NUMBER: US 60/429,609
; PRIOR FILING DATE: 2002-11-27
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 390
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-437-427-6
```

```
Query Match 15.8%; Score 307.5; DB 15; Length 390;
Best Local Similarity 25.3%; Pred. No. 3.1e-16;
Matches 95; Conservative 63; Mismatches 151; Indels 67; Gaps 13;
QY 24 FAFSVRWSDGSDTFVRRSWDEFRQLKTLKETFPVEAGLLRRSDRVLPKLLDAPLLGRVG 83
Db 26 YMFVKWQDLSEKVVYRRFTIYEFHKLTKEMFPIEAGAINPENRIIPH-LPAPKWFQDG 84
QY 84 RTSRGLARQLLETYSRRLATAERVARSPITITGFFAPQPLDLEPALPPGSRV-----I 137
Db 85 RAAEN--RQGLTEYCVSTLMSLPTKISCPHLLDFKVRPDDLK--LPTDNQTKKPEYTL 140
QY 138 LPTPEEQLSRAAGRLSIHSLAQSLRCLQPCQTODTRDRPFQAQAESLDVLLRHPGSM 197
Db 141 MPKDGKSTATDITGPILL-----QTYRAIADY--EKTSGSEWALSTGDVVEVKESEGW 193
QY 198 WLVENEDQTAFFAPYLEEA-APGQREGGPGSLGSSGPFQFCASRAYESSRADELSPVAG 256
Db 194 WFCQMKAKR-GWIPASFLEPLDSPDETDPENY--AGEPYVAIKAYTAVBGDEVSLLEG 250
QY 257 ARVRVLETSDRGWLCRYGDRAGLLPAVLLRPEG-----LGA----- 293
Db 251 EAVEVIHKLDDGWVIRKDDVTGYFFPSMYLQSGQDVSAQORQIKRGAPRRSSIRNAHS 310
QY 294 -----LLSGTGFRRGD-----DPAGEARGFP-----EFSQATAPPTVTPRPS 331
Db 311 IHQSRKRLSDAYRNSVRFLOQRROARPGSPGSPLEBERQTRSKQPAPVPPRS 370
QY 332 PGAIQSRCTVTTRAL 347
Db 371 ADLILNRCSESTKRKL 386
```

RESULT 9

```
US-10-767-341-4
; Sequence 4, Application US/10767341
; Publication No. US20040132084A1
; GENERAL INFORMATION:
; APPLICANT: SHAO, Wei et al.
; TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001198DIV-II
; CURRENT APPLICATION NUMBER: US/10/767,341
; CURRENT FILING DATE: 2004-01-30
; PRIOR APPLICATION NUMBER: 09/820,005
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 10/109,856
; PRIOR FILING DATE: 2002-04-01
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 390
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-767-341-4
```

```
Query Match 15.8%; Score 307.5; DB 16; Length 390;
```



```
US-10-437-427-2
; Sequence 2, Application US/10437427
; Publication No. US2004000901A1
; GENERAL INFORMATION:
; APPLICANT: Rikard Olofsson
; TITLE OF INVENTION: Autoimmune Conditions and NADPH Oxidase
; TITLE OF INVENTION: Defects
; FILE REFERENCE: 11145-024001
; CURRENT APPLICATION NUMBER: US/10/437,427
; CURRENT FILING DATE: 2003-05-13
; PRIOR APPLICATION NUMBER: US 60/380,904
; PRIOR FILING DATE: 2002-05-13
; PRIOR APPLICATION NUMBER: US 60/429,609
; PRIOR FILING DATE: 2002-11-27
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 389
; TYPE: PRT
; ORGANISM: Rattus norvegicus
; ORGANISM: Rattus norvegicus
US-10-437-427-2

Query Match
Best Local Similarity 15.7%; Score 306; DB 15; Length 389;
Matches 101; Conservative 55; Mismatches 155; Indels 62; Gaps 14;

QY 24 FAFSVRWSGSDTFFVRRSWDEFRQLKTKLTFPVEAGLLRRSDRVLPKLLDAPLLGRVG 83
D6 YMFVWKQDLSKVVYRFTIYEFHKLKEMFPIEAGIHTENRVIPIH-LPAPRWYDQ 84
QY 84 RTSRGLARLQLLETYSRLLATAERVARSPITGFAFQPLDLEPALPGSRVILPTPEE 143
D6 85 RAAE--SRQGLTEYFNSLMGLPMKISRCPHLLNFKFVRPDDLK--LPNDSQVKKP---E 137
QY 144 QLSRAAGRLSIHSLA-----QSLCLQPFCTQDTRDRPFAQAQESLDVLLRHPSGWL 199
D6 138 TYLTAKGKNVADINPIIQTYYRAIDY--EKSKTEMTVATGDVDDVVEKSSGNWF 195
QY 200 VENEDRQTAWFPAPYLEEA-AFGQREGGPGSLGSGQPCFASRAYESSRADELSVPAGAR 258
D6 196 CQMKTKR-GWIPASYLEPLDSDPEAEDDPNY--AGEPVYTIKAYAAVEEDSVLSGEA 252
QY 259 VRVLETSRGRWLCRYGDRAGLLPAVLRPEG-----LGA----- 293
D6 253 IEVIHKLLDGMVVRKGDITGYFPMSYLOKAGEITQAQRQIRSRGAPPRKSTIRNAQSI 312
QY 294 -----LLSGTGRGGD-----DPAGEARGFPE---PSQATA-PPPTVPTPRSPGA 334
D6 313 HORSKRKLSQDYRRNSVRFLQORRRPAPPGQSPGSPLEEBERQTSKQTPQAPVPPRPSDL 372
QY 335 IQSRCTCTVTRRAL 347
D6 373 ILHRCSTESTKRKL 385

RESULT 13
US-10-109-856-2
; Sequence 2, Application US/10109856
; Publication No. US20030166185A1
; GENERAL INFORMATION:
; APPLICANT: SHAO, Wei et al.
; TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001198DIV
; CURRENT APPLICATION NUMBER: US/10/109,856
; CURRENT FILING DATE: 2002-04-01
; PRIOR APPLICATION NUMBER: 09/820,005
; PRIOR FILING DATE: 2001-03-29
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 386
; TYPE: PRT
; ORGANISM: Homo sapiens
; ORGANISM: Homo sapiens
US-10-109-856-2

Query Match
Best Local Similarity 14.3%; Score 278.5; DB 16; Length 386;
Matches 93; Conservative 62; Mismatches 150; Indels 71; Gaps 14;

QY 24 FAFSVRWSGSDTFFVRRSWDEFRQLKTKLTFPVEAGLLRRSDRVLPKLLDAPLLGRVG 83
D6 26 YMFVWKQDLSKVVYRFTIYEFHKLKEMFPIEAGIHTENRVIPIH-LPAPRWYDQ 84
QY 84 RTSRGLARLQLLETYSRLLATAERVARSPITGFAFQPLDLEPALPGSRVILPTPEE 137
D6 85 RAAE--RQGLTEYCSLTMSLPTKISRCPHLLDFFKVRPDDLK--LPTDNQTKKPETYL 140
QY 138 LPTPEEQPLSRAAGRLSIHSLAQSRLCLQPFCTQDTRDRPFAQAQESLDVLLRHPSGW 197
D6 141 MPKDGKSTATDITGPILL-----QTYRAINY--EKTSGSEMASTGDVVEVEKSESGW 193
QY 198 WLVENEDRQTAWFPAPYLEEA-AFGQREGGPGSLGSGQPCFASRAYESSRADELSVPAG 256
D6 194 WFCQMKAKR-GWIPASYLEPLDSDPEAEDDPNY--AGEPVYTIKAYTAVEGDEVSLLEG 250
QY 257 ARVRVLETSRGRWLCRYGDRAGLLPAVLRPEG-----LGA----- 293
D6 251 EAVEVIHKLLDGM-----KDDVTGYFPMSYLOKSGQVDSQAQRQIKRGAPPRRRSSIRNAHS 306
QY 294 -----LLSGTGRGGD-----DPAGEARGFPE---BPSQATAPPPTVPTPRPS 331
D6 307 IHORSKRKLSQDYRRNSVRFLQORRRPAPPGQSPGSPLEEBERQTSKQTPQAPVPPRPS 366
QY 332 PGAIQSRCTCTVTRRAL 347
D6 367 ADLILNRCSESTKRKL 382

RESULT 14
US-10-767-341-2
; Sequence 2, Application US/10767341
; Publication No. US20040132084A1
; GENERAL INFORMATION:
; APPLICANT: SHAO, Wei et al.
; TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001198DIV-II
; CURRENT APPLICATION NUMBER: US/10/767,341
; CURRENT FILING DATE: 2004-01-30
; PRIOR APPLICATION NUMBER: 09/820,005
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 10/109,856
; PRIOR FILING DATE: 2002-04-01
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 386
; TYPE: PRT
; ORGANISM: Homo sapiens
; ORGANISM: Homo sapiens
US-10-767-341-2

Query Match
Best Local Similarity 14.3%; Score 278.5; DB 16; Length 386;
Matches 93; Conservative 62; Mismatches 150; Indels 71; Gaps 14;

QY 24 FAFSVRWSGSDTFFVRRSWDEFRQLKTKLTFPVEAGLLRRSDRVLPKLLDAPLLGRVG 83
D6 26 YMFVWKQDLSKVVYRFTIYEFHKLKEMFPIEAGIHTENRVIPIH-LPAPRWYDQ 84
QY 84 RTSRGLARLQLLETYSRLLATAERVARSPITGFAFQPLDLEPALPGSRVILPTPEE 137
D6 85 RAAE--RQGLTEYCSLTMSLPTKISRCPHLLDFFKVRPDDLK--LPTDNQTKKPETYL 140
QY 138 LPTPEEQPLSRAAGRLSIHSLAQSRLCLQPFCTQDTRDRPFAQAQESLDVLLRHPSGW 197
D6 141 MPKDGKSTATDITGPILL-----QTYRAINY--EKTSGSEMASTGDVVEVEKSESGW 193
QY 198 WLVENEDRQTAWFPAPYLEEA-AFGQREGGPGSLGSGQPCFASRAYESSRADELSVPAG 256
D6 194 WFCQMKAKR-GWIPASYLEPLDSDPEAEDDPNY--AGEPVYTIKAYTAVEGDEVSLLEG 250
QY 257 ARVRVLETSRGRWLCRYGDRAGLLPAVLRPEG-----LGA----- 293
D6 251 EAVEVIHKLLDGM-----KDDVTGYFPMSYLOKSGQVDSQAQRQIKRGAPPRRRSSIRNAHS 306
QY 294 -----LLSGTGRGGD-----DPAGEARGFPE---BPSQATAPPPTVPTPRPS 331
D6 307 IHORSKRKLSQDYRRNSVRFLQORRRPAPPGQSPGSPLEEBERQTSKQTPQAPVPPRPS 366
QY 332 PGAIQSRCTCTVTRRAL 347
D6 367 ADLILNRCSESTKRKL 382
```

```
US-10-437-427-2
; Sequence 2, Application US/10437427
; Publication No. US2004000901A1
; GENERAL INFORMATION:
; APPLICANT: Rikard Olofsson
; TITLE OF INVENTION: Autoimmune Conditions and NADPH Oxidase
; TITLE OF INVENTION: Defects
; FILE REFERENCE: 11145-024001
; CURRENT APPLICATION NUMBER: US/10/437,427
; CURRENT FILING DATE: 2003-05-13
; PRIOR APPLICATION NUMBER: US 60/380,904
; PRIOR FILING DATE: 2002-05-13
; PRIOR APPLICATION NUMBER: US 60/429,609
; PRIOR FILING DATE: 2002-11-27
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 389
; TYPE: PRT
; ORGANISM: Rattus norvegicus
; ORGANISM: Rattus norvegicus
US-10-437-427-2
```

```
Query Match
Best Local Similarity 15.7%; Score 306; DB 15; Length 389;
Matches 101; Conservative 55; Mismatches 155; Indels 62; Gaps 14;

QY 24 FAFSVRWSGSDTFFVRRSWDEFRQLKTKLTFPVEAGLLRRSDRVLPKLLDAPLLGRVG 83
D6 26 YMFVWKQDLSKVVYRFTIYEFHKLKEMFPIEAGIHTENRVIPIH-LPAPRWYDQ 84
QY 84 RTSRGLARLQLLETYSRLLATAERVARSPITGFAFQPLDLEPALPGSRVILPTPEE 143
D6 85 RAAE--SRQGLTEYFNSLMGLPMKISRCPHLLNFKFVRPDDLK--LPNDSQVKKP---E 137
QY 144 QLSRAAGRLSIHSLA-----QSLCLQPFCTQDTRDRPFAQAQESLDVLLRHPSGWL 199
D6 138 TYLTAKGKNVADINPIIQTYYRAIDY--EKSKTEMTVATGDVDDVVEKSSGNWF 195
QY 200 VENEDRQTAWFPAPYLEEA-AFGQREGGPGSLGSGQPCFASRAYESSRADELSVPAGAR 258
D6 196 CQMKTKR-GWIPASYLEPLDSDPEAEDDPNY--AGEPVYTIKAYAAVEEDSVLSGEA 252
QY 259 VRVLETSRGRWLCRYGDRAGLLPAVLRPEG-----LGA----- 293
D6 253 IEVIHKLLDGMVVRKGDITGYFPMSYLOKAGEITQAQRQIRSRGAPPRKSTIRNAQSI 312
QY 294 -----LLSGTGRGGD-----DPAGEARGFPE---PSQATA-PPPTVPTPRSPGA 334
D6 313 HORSKRKLSQDYRRNSVRFLQORRRPAPPGQSPGSPLEEBERQTSKQTPQAPVPPRPSDL 372
QY 335 IQSRCTCTVTRRAL 347
D6 373 ILHRCSTESTKRKL 385
```

```
RESULT 13
US-10-109-856-2
; Sequence 2, Application US/10109856
; Publication No. US20030166185A1
; GENERAL INFORMATION:
; APPLICANT: SHAO, Wei et al.
; TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001198DIV
; CURRENT APPLICATION NUMBER: US/10/109,856
; CURRENT FILING DATE: 2002-04-01
; PRIOR APPLICATION NUMBER: 09/820,005
; PRIOR FILING DATE: 2001-03-29
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
```


This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 28, 2005, 07:26:27 ; Search time 40 Seconds
(without alignments)
892.410 Million cell updates/sec

Title: US-10-621-113-4
Perfect score: 1950
Sequence: 1 MGPRYPVSVQGAALVQIKR.....RRQGRPRGCVDSVPHTTQ 371
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79:.*
1: pirl:.*
2: pirl:.*
3: pirl:.*
4: pirl:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	307.5	15.8	390	1 A39249	neutrophil cytosol
2	300.5	15.4	388	2 I54525	leukemia-related p
3	161	8.3	940	2 T00056	hypothetical prote
4	151	7.7	1270	2 T09194	adaptor protein in
5	139.5	7.2	560	1 WFBOM	mullerian inhibiti
6	136.5	7.0	1168	1 MWAXIC	myosin heavy chain
7	128.5	6.6	377	2 S08636	nck protein - huma
8	128	6.6	1249	2 A56511	myosin I myoA - Em
9	125	6.4	1094	2 T13053	dynamain associated
10	123	6.3	1011	2 T13055	hypothetical prote
11	122.5	6.3	290	2 T42526	probable bifunctio
12	121	6.2	463	2 T34841	hypothetical prote
13	120.5	6.2	1097	2 T15104	saframycin Mx1 syn
14	118.5	6.1	2605	2 T18552	mixed-lineage prot
15	118	6.1	847	1 A53800	hypothetical prote
16	117	6.0	443	2 T27877	hypothetical prote
17	116.5	6.0	780	2 T00366	probable large Pro
18	116	5.9	1366	2 T35985	Sufi protein (impo
19	115	5.9	470	2 AD0888	protein-tyrosine k
20	115	5.9	512	2 I49552	F02569_2 protein [
21	115	5.9	512	2 E59437	CRKL protein - hum
22	114.5	5.9	303	2 S41754	LckBp1 protein - m
23	114.5	5.9	486	2 I49760	hypothetical prote
24	114	5.8	364	2 T35353	probable ATP-depen
25	113.5	5.8	639	2 D83591	C-Crk - chicken
26	112	5.7	305	1 A49011	SH2/SH3 adaptor pr
27	111.5	5.7	303	2 S58352	C-Crk - mouse
28	111	5.7	304	2 I58394	B120 protein - hum
29	110.5	5.7	1142	2 T00022	

ALIGNMENTS

RESULT 1

A39249
neutrophil cytosol factor 1 - human
N;Alternate names: 47K autosomal chronic granulomatous disease protein; multicomponent m
C;Species: Homo sapiens (man)
C;Date: 30-Jun-1993 #sequence revision 30-Jun-1993 #text change 09-Jul-2004
C;Accession: A35926; A39249; A54067; I59190; A32762; A41385
R;Rodaway, A.R.F.; Teahan, C.G.; Casimir, C.M.; Segal, A.W.; Bentley, D.L.
Mol. Cell. Biol. 10, 5388-5396, 1990
A;Title: Characterization of the 47-kilodalton autosomal chronic granulomatous disease p
A;Reference number: A35926; MUID:90377229; PMID:2398896
A;Accession: A35926
A;Molecule type: mRNA
A;Residues: 1-390 <RD>
A;Cross-references: UNIPROT:PI4598; GB:M55067; GB:M38755; NID:gl89050; PIDN:AA59901.1; I
R;Volpp, B.D.; Nauseef, W.M.; Donelson, J.E.; Moser, D.R.; Clark, R.A.
Proc. Natl. Acad. Sci. U.S.A. 86, 9563, 1989
A;Reference number: A39249
A;Contents: erratum
A;Accession: A39249
A;Molecule type: mRNA
A;Residues: 1-390 <VOL>
A;Cross-references: GB:M26193; NID:gl89107; PIDN:AA57209.1; PID:gl89108
R;Volpp, B.D.; Nauseef, W.M.; Donelson, J.E.; Moser, D.R.; Clark, R.A.
Proc. Natl. Acad. Sci. U.S.A. 86, 7195-7199, 1989
A;Title: Cloning of the cDNA and functional expression of the 47-kilodalton cytosolic com
A;Reference number: A41385; MUID:89386707; PMID:2550933
A;Contents: annotation
A;Note: the sequence reported has been extensively revised in reference A39249
R;Lomax, K.J.; Leto, T.L.; Nunoi, H.; Gallin, J.I.; Malech, H.L.
Science 245, 409-412, 1989
A;Title: Recombinant 47-kilodalton cytosol factor restores NADPH oxidase in chronic gran
A;Reference number: A32762; MUID:89332501; PMID:2547247
A;Contents: annotation
A;Note: the sequence reported has been extensively revised and now agrees with that show
R;Finan, P.; Shimizu, Y.; Gout, I.; Hsuan, J.; Truong, O.; Butcher, C.; Bennett, P.; Wat
J. Biol. Chem. 269, 13752-13755, 1994
A;Title: An SH3 domain and proline-rich sequence mediate an interaction between two comp
A;Reference number: A54067; MUID:94245680; PMID:8188650
A;Accession: A54067
A;Molecule type: protein
A;Residues: 8-16;44-52;71-77 <FIN>
A;Experimental source: differentiated HL-60 cells
R;Casimir, C.M.; Bu-Gharios, H.N.; Rodaway, A.R.F.; Bentley, D.L.; Rowe, P.; Segal, A.W.
Proc. Natl. Acad. Sci. U.S.A. 88, 2753-2757, 1991
A;Title: Autosomal recessive chronic granulomatous disease caused by deletion at a dinucl
A;Reference number: I59190; MUID:91187870; PMID:2011585
A;Accession: I59190
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 14-24 <CAS>
A;Cross-references: GB:M60941; NID:gl899948; PIDN:AAA60086.1; PID:gl899949

hypothetical prote
N-methyl-D-asparta
hypothetical prote
probable membrane
serine/threonine p
hypothetical prote
probable serine/th
DNA helicase RecG
hypothetical prote
rho-GTPase-activat
hypothetical prote
mullerian inhibiti
inositol 1,4,5-tri
nascent polypeptid
translation initia
BOLFI protein - hu

30 110 5.6 603 2 T43489
31 110 5.6 1356 1 C45219
32 109.5 5.6 878 2 T17245
33 108.5 5.6 797 2 T34966
34 108.5 5.6 1032 2 D83637
35 108 5.5 384 2 H70580
36 107.5 5.5 720 2 T36293
37 107.5 5.5 784 2 H75338
38 107 5.5 408 2 T42650
39 107 5.5 946 2 I38100
40 106.5 5.5 1386 2 T00257
41 106 5.4 575 1 WFBOM
42 105.5 5.4 946 2 JC7810
43 105.5 5.4 2187 2 T30826
44 104.5 5.4 1037 2 B87254
45 104.5 5.4 1239 1 QQB810

A>Note: a GT dinucleotide deletion at a GT-GT repeat causes a frameshift after residue 2
C:Comment: This protein is required for activation of the latent NADPH oxidase, which is
ous disease.

C:Genetics:
A:Gene: GDB:NCF1
A:Cross-references: GDB:120222; OMIM:233700
A:Map position: 7q11.23-7q11.23
A:Introns: 24/3
A>Note: the list of introns is incomplete
C:Superfamily: neutrophil cytosol factor 1; SH3 homology
C:Keywords: cytosol; neutrophil
F:163-210/Domain: SH3 homology <SH31>
F:233-280/Domain: SH3 homology <SH32>

Query Match 15.8%; Score 307.5; DB 1; Length 390;
Best Local Similarity 25.3%; Pred. No. 2.1e-14;
Matches 95; Conservative 63; Mismatches 151; Indels 67; Gaps 13;

```
QY 24 FAFSVRWSGSDTFVRRSWDFRQLKKTLETFPPVEAGLLRRSDRVLPKLLDAPLLGRVG 83
DB 26 YMFVWKQDLSEKVVYRRFTIYEFHKLKEMFPPIEAGAINPENRIIPH-LPAPKWFQDQ 84
QY 84 RTSRGLARLQLLETYSRRLATAERVARSPITITGFFAQPQLDLEPALPGSRV-----I 137
DB 85 RAAEN--RQGLTLCSTLSLPTKISRCPLHLLDFKVRPDDLK--LPDNTQTKKPEYVL 140
QY 138 LPTPEQPLSRAAGRLSHLSLEAQLRCLQPCQTQDTRDRPPQAQAESLQVLRHPSGW 197
DB 141 MPKDGKSTATDITGPILL-----QTYRAIADY--BKTSGSEMALSTGDVVEVESGSGW 193
QY 198 WLVENEDRTQAFWPAPYLEEA-APGQREGGSLGSSGPPQFCASRAYESSRADELSPVAG 256
DB 194 WFCQMKAR-GWIPASFLPLDSPDETEPEPNY--AGEPYVAIKAYTAVEGDEVSLLEG 250
QY 257 ARVRYLETSDRGWMLCRYGDRAGLLPAVLLRPEG-----LGA----- 293
DB 251 EAVEVHKLLDGMWVRKDDVTGYFSPMYLQSGQDVQAQRIKRGAPPRSSIRNAHS 310
QY 294 -----LLSGTGFRRGGD-----DPAGEARGPP-----EPGQATAPPTVTPRPS 331
DB 311 IHQSRKRLSQDQAYRRNSVRFLLQRRRQARPGQSPGSPLEERQTORSKPQPAVPPRPS 370
QY 332 PGAIQSRCTVTTRAL 347
DB 371 ADLIILNRCSESTKRKL 386
```

RESULT 2

leukemia-related protein - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 16-Jul-1999
C:Accession: I54525
R:Jackson, S.H.; Malech, H.L.; Kozak, C.A.; Lomax, K.J.; Gallin, J.I.; Holland, S.M.
Immunogenetics 39, 272-275, 1994
A:Title: Cloning and functional expression of the mouse homologue of p47phox.
A:Reference number: I54525; MUID:94164697; PMID:8119734
A:Accession: I54525
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-388 <RES>
A:Cross-references: GB:L11455; NID:g309422; PIDN:AAA50469.1; PID:g557868
C:Superfamily: neutrophil cytosol factor 1; SH3 homology
F:161-208/Domain: SH3 homology <SH31>

Query Match 15.4%; Score 300.5; DB 2; Length 388;
Best Local Similarity 25.7%; Pred. No. 6.5e-14;
Matches 96; Conservative 55; Mismatches 160; Indels 63; Gaps 11;

```
QY 24 FAFSVRWSGSDTFVRRSWDFRQLKKTLETFPPVEAGLLRRSDRVLPKLLDAPLLGRVG 83
DB 24 YMFVWKQDLSEKVVYRRFTIYEFHKLKEMFPPIEAGIHTENRVIIPH-LPAPRWFDQ 82
```

```
QY 84 RTSRGLARLQLLETYSRRLATAERVARSPITITGFFAQPQLDLEPALPGSRVILPTPEE 143
DB 83 RAAE--SRQGLTLYFENGLMGLPVKISRCPLHLLDFKVRPDDLK--LPTDSQAKP---E 135
QY 144 QPLSRAAGRLSHLSLEA---QSLRCLQPFCTQDTRDRPPQAQAESLDVLRHPSGWWL 199
DB 136 TVLVPKDGNVADITGPILLQTHRAIADY--EKSSTGTEMTVATGDVVVDVVEKESGWWF 193
QY 200 VENEDRTQAFWPAPYLEEA-APGQREGGSLGSSGPPQFCASRAYESSRADELSPAGAR 258
DB 194 CQMKTKR-GWVPASYLEPLDSPDEAEDPPNY--AGEPYVTIKAYAAVEDEMSLSEGEA 250
QY 259 VRVLETSDRGWMLCRYGDRAGLLPAVLLRPPGLGALLSGTGRGDDPAGEA----- 310
DB 251 IEVHKLLDGMWVRKDDITGYFFSWMYLOKAGEBITQQRQIRGRCAPPRSTIRNAOSI 310
QY 311 --RGFPPEPSQAT----- 333
DB 311 HQSRKRLSQDQAYRRNSVRFLLQRRRPGRPRAASTDGTGKNDPSTPRVKPQPAVPPRPSD 370
QY 334 AIQSRCTVTTRAL 347
DB 371 LILHRCSTESTKRKL 384
```

RESULT 3

hypothetical protein KIAA0418 - human
C:Species: Homo sapiens (man)
C:Date: 22-Jan-1999 #sequence_revision 22-Jan-1999 #text_change 09-Jul-2004
C:Accession: T00056
R:Ishikawa, K.; Nagase, T.; Nakajima, D.; Seki, N.; Ohira, M.; Miyajima, N.; Tanaka, A.;
submitted to the EMBL Data Library, October 1997
A:Description: Prediction of the coding sequences of unidentified human genes. VIII. The
A:Reference number: Z14080
A:Accession: T00056
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-940 <ISH>
A:Cross-references: UNIPROT:O43302; EMBL:AB007878; NID:d1179754; PIDN:BAA24848.1; PID:d11
A:Experimental source: brain
C:Genetics:
A>Note: KIAA0418

Query Match 8.3%; Score 161; DB 2; Length 940;
Best Local Similarity 24.4%; Pred. No. 0.0012;
Matches 53; Conservative 23; Mismatches 73; Indels 68; Gaps 8;

```
QY 182 QAQESLDVLLRHPGSMWLVENEDRTQAFWPAPYLEEAAPGQREGGSLGSSGPPQ--FCA 239
DB 23 QAGEVVDVIEKNESGMWPFVSTSEQ-GWVPATYL-EAQNGTRDDSDINTSKTGESEKVT 80
QY 240 SRAYESSRADELSPAGARVVRVLETSDRGWMLCRYGDRAGLLPAVLLRP----- 288
DB 81 VQPYTSQSKDIBIGFEKGVTVVEIRKLEGGWVYRVLGEGWAPASLYLKAKKDDLPTRKKN 140
QY 289 -----ELGALLSGTGF-----RGGD---DPA-----GEARG 312
DB 141 LAGPVEIIGNIMEISNLNKKASGDKETPTPAGEGHEAPIAKKEISLPILCNASNGSAVG 200
QY 313 FPE-----PSQATAPPTVTPRPS 332
DB 201 VPDRTVSRLAQSPAVARIAIAPQAISSPNLRTTRPPP 237
```

RESULT 4

T09194
adaptor protein intersectin - African clawed frog
C:Species: Xenopus laevis (African clawed frog)
C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 09-Jul-2004
C:Accession: T09194
R:Yamabhai, M.; Hoffman, N.G.; Hardison, N.L.; McPherson, P.S.; Castagnoli, L.; Cesareni,
J. Biol. Chem. 273, 31401-31407, 1998

A;Title: Intersectin, a novel adaptor protein with two eps15 homology and five src homologous domain regions
A;Reference number: Z16605; MUID:99030416; PMID:9813051
A;Accession: T09194
A;Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A;Residues: 1-1270 <VAM>
A;Cross-references: UNIPROT:O42287; EMBL:AF032118; NID:g2642624; PIDN:AAC73068.1; PID:g2642624
A;Experimental source: cell type oocyte
C:Function:
A;Description: Involved in endocytosis
C:Keywords: endocytosis

Query Match 7.7%; Score 151; DB 2; Length 1270;
Best Local Similarity 27.9%; Pred.No.0.0085;
Matches 61; Conservative 18; Mismatches 80; Indels 60; Gaps 12;

QY 179 FOAQAESLD-----VLLR--HPSGWLVVE----NEDROTAMPFPAPYLEEAAPGGGRE- 225
DB :.:|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.:|.:.|.:|.:.
QY 226 -GPGSLSSGPQC---ASRAYESSRADELSPAGARVRVLETSDRGWMLCRYDGRAGLL 281
DB :.:|.|.|.|.|.|.|.|.|.|.:|.:.|.:|.:.|.:|.:.|.:|.:.
QY 282 PA-VLLRPEGLALLSGTGFRCDDPAGARGFPPSQNTAPPVT--PTRSPG----- 333
DB :.:|.|.|.|.|.|.|.|.|.|.:|.:.|.:|.:.|.:|.:.|.:|.:.
QY 1189 PSNYVKLTDT----MDPSQQFRLGVPAG---GTP-----ATGDRPTFLFFRDGPSLLPN 1237
DB :.:|.|.|.|.|.|.|.|.|.|.:|.:.|.:|.:.|.:|.:.|.:|.:.
QY 334 -----AIQSRTCCTVTRALERPRRGRCGCVD 362
DB :.:|.|.|.|.|.|.|.|.|.|.:|.:.|.:|.:.|.:|.:.|.:|.:.
QY 1238 AFQAPLSVVMKFCFTA-----PFPCPD 1262
DB :.:|.|.|.|.|.|.|.|.|.|.:|.:.|.:|.:.|.:|.:.|.:|.:.

RESULT 5
WFHUM
mullerian inhibiting factor precursor - human
N;Alternate names: anti-Müllerian hormone; mullerian inhibiting substance (MIS)
C:Species: Homo sapiens (man)
C:CDate: 13-Aug-1986 #sequence_revision 13-Aug-1986 #text_change 09-Jul-2004
C;Accession: A01397
R;Cate: R.L.; Mattaliano, R.J.; Hession, C.; Tizard, R.; Farber, N.M.; Cheung, A.; Ninfa,
an, K.L.; Ragin, R.C.; Mangano, T.F.; MacLaughlin, D.T.; Donahoe, P.K.
Cell 45, 685-698, 1986
A;Title: Isolation of the bovine and human genes for Mullerian inhibiting substance and
antiserum against it
A;Reference number: A90879; MUID:86218082; PMID:3754790
A;Accession: A01397
A:Molecule type: DNA
A;Residues: 1-560 <UNT>
A;Cross-references: UNIPROT:P03971; GB:X03474; NID:g188560; PIDN:AAA98805.1; PID:g386953
C;Comment: Although it does not compete with EGF for receptor binding sites, MIS can inh
C;Comment: For anti-Müllerian hormone type II receptor, see PIR:JC435.
C;Genetics:
A;Gene: GDB:NMH
A;Cross-references: GDB:118996; OMIM:261550; OMIM:600957
A;Map position: 19p13.3-19p13.3
A;Introns: 138/1; 185/3; 222/1; 275/2
C;Superfamily: inhibit
C;Keywords: cytotoxin; glycoprotein; gonadal differentiation; hormone; testis
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-25/Domain: propeptide #status predicted <PRO>
F:26-560/Product: mullerian inhibiting factor #status predicted <MAT>
F:64,329/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:462-526,488-557,492-559/Disulfide bonds: #status predicted
F:525/Disulfide bonds: interchain #status predicted

Query Match 7.2%; Score 139.5; DB 1; Length 560;
Best Local Similarity 28.6%; Pred.No.0.022;
Matches 93; Conservative 20; Mismatches 147; Indels 65; Gaps 18;

QY 89 LARLOLETYSRRLLATAERVARSPFI--TGFPAPQLDLPALP--PGSRVL----PTP 141
DB :.:|.|.|.|.|.|.|.|.|.|.:|.:.|.:|.:.|.:|.:.|.:|.:.
DB 6 LTSLALVLSALGALLGTALRAEPVAVGSLFREDLOWPGIPOEPLCLVALGDGSNG 65

QY 229 S-----LGSSGPOFCAS-----RA---YESSRADELSVPAGARVRVLETSDRGWMLCR 273
Db 958 SFGGRPSFGSPPAAASAPGPEQARALYDFAENDELTFNEGAVVTWINKSNPDWGE 1017
QY 274 YGDRAGLLPAVLL-----RP-----EGLCALLSGTG---FRGGDDPA 307
Db 1018 LNGQGVFPASIVELIPRAAAPGSPGPRAPPGKSGRAAPMGFGPMRGSGPAPG 1077
QY 308 GEAR-GFPEPSQATAPPVTRPSPGAIQSRCCCTVTRRALERRPRGRP 357
Db 1078 GPRGGAPPGAGRAGPGGRCMPAPGG-----AAPRGAP 1114
RESULT 7
S08636
nck protein - human
N:Alternate names: src-related protein
C:Species: Homo sapiens (man)
C:Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 09-Jul-2004
C:Accession: S08636
R:Lehmann, J.M.; Riethmuller, G.; Johnson, J.P.
Nucleic Acids Res. 18, 1048, 1990
A:Title: Nck, a melanoma cDNA encoding a cytoplasmic protein consisting of the src homol
A:Reference number: S08636; MUID:90192089; PMID:2107526
A:Accession: S08636
A:Molecule type: mRNA
A:Residues: 1-377 <LEH>
A:Cross-references: UNIPROT:P16333; EMBL:X17576; NID:g35014; PIDN:CAA35599.1; PID:g35015
F:9-56/Domain: SH3 homology <SH31>
F:113-160/Domain: SH3 homology <SH32>
F:197-247/Domain: SH3 homology <SH33>
F:282-371/Domain: SH2 homology <SH2>

Query Match 6.6%; Score 128.5; DB 2; Length 377;
Best Local Similarity 24.6%; Pred. No. 0.083;
Matches 44; Conservative 20; Mismatches 52; Indels 63; Gaps 6;
QY 179 FOAQAESLDV-----LLRHPSGWLVENEDROTANFPAPYLE-----E 217
Db 13 YVAQOEQLDIKKERLMLDDSKSWVRVNSMNTGFVPSNYVERKNSARKASIVKNLK 72
QY 218 AAPGGR-EGGPSLGGSPQFCASRA-----YESSRADELSV 253
Db 73 DTLGIGKVRKSPVDS-----ASPADDSFVDPGRLYDLNMPAVKYNWAEDELSL 127
QY 254 PAGARVRVLETSDRGWMLCRYGDRAGLLPAVLLRPEGLGALLSGTGFRGGDDPAGEARG 312
Db 128 IKGTVIWEKSCDGMWRSYNGVQGWFPNSVYTEE-----GDSPLGDHVG 173

RESULT 8
A56511
myosin I myoA - Emericella nidulans
C:Species: Emericella nidulans, Aspergillus nidulans
C:Date: 21-Jul-1995 #sequence_revision 28-Jul-1995 #text_change 09-Jul-2004
C:Accession: A56511
J. Cell Biol. 128, 577-587, 1995
A:Title: myoA of Aspergillus nidulans encodes an essential myosin I required for secreti
A:Reference number: A56511; MUID:95164560; PMID:7860631
A:Accession: A56511
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1249 <MCG>
A:Cross-references: UNIPROT:Q00647; GB:U12427; NID:G525321; PIDN:AAA67877.1; PID:G525322
C:Genetics:
A:Gene: myoA
C:Superfamily: protozoan myosin heavy chain IB; myosin motor domain homology; SH3 homolo
C:Keywords: nucleotide binding; P-loop
F:53-716/Domain: myosin motor domain homology <MMOT>
F:143-150/Region: nucleotide-binding motif A (P-loop)
F:1081-1130/Domain: SH3 homology <SH3>

Query Match 6.6%; Score 128; DB 2; Length 1249;
Best Local Similarity 23.4%; Pred. No. 0.35;
Matches 62; Conservative 25; Mismatches 92; Indels 86; Gaps 12;
QY 128 PALPFGSRVILPTPE-----EQPLSRAAGRLSIHSLEA-----QSURC 165
Db 970 PGEPPNS-VSKPTPRGQVAARPVTK--GKLLRLAVQAVARPNNWLPDLYQSVGLYHSPL 1026
QY 166 LQPCTQDTRDRPFOAQAESLDVLLRHPSGWLVENEDROTANFPAPYLEEAPGQGRE 225
Db 1027 KQPERNRHQRDPDFLNQWP-----LQHPI-----HVLHLLPPQGH 1063
QY 226 GGPSL-----GSSGPOFC-ASRAYESSRADELSVPAGARVRVLETSDRGWMLCRYGDRA- 278
Db 1064 --PRLLPRPPAAAGPKAKALYDFSSDNNNGMLSIAGQIVEISKEGNGWMLCNLETSA 1121
QY 279 -GLLPAVLLRPEGLGALLSGTGFRGGDDPAGEARGFPEPSQATAPPVTRPSPGAIQS 337
Db 1122 QGWTPEAVL-----EEQVAPTFKPAPPPPPVPVAPRASPAPVNG 1159
QY 338 RCTVTTRAL-----ERRRRQR 356
Db 1160 SAAVAATAKAKAAPPPPAKRPENMAGR 1184
RESULT 9
TI3053
dynamain associated protein isoform Dapl60-1 - fruit fly (Drosophila melanogaster)
C:Species: Drosophila melanogaster
C:Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 09-Jul-2004
C:Accession: TI3053
R:Roos, J.; Kelly, R.B.
J. Biol. Chem. 273, 19108-19119, 1998
A:Title: Dapl60, a neural-specific Eps15 homology and multiple SH3 domain-containing prot
A:Reference number: Z17594; MUID:98334647; PMID:9668096
A:Accession: TI3053
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1094 <ROO>
A:Cross-references: UNIPROT:O61618; EMBL:AF0533957; NID:g2984714; PID:g2984715; PIDN:AAAC33
C:Genetics:
A:Gene: Dapl60
A:Cross-references: FlyBase:FBgn0023388
Query Match 6.4%; Score 125; DB 2; Length 1094;
Best Local Similarity 23.9%; Pred. No. 0.49;
Matches 55; Conservative 33; Mismatches 86; Indels 56; Gaps 10;
QY 106 AERVARPTITGFPAPQPLDEPA-LPPGSRVILPTPE---QPLSRAA---GRLSIHS 157
Db 881 SEQPISSFGVGAERAEHDLTEVSQINTQKTSQSEPAESYSRPMRSRTSSMTPGMKRKS 940
QY 158 LEAQSRLQPFCTQDTRDRPFOAQAESLD-----VLLRH---PSGWM-----LVENEDR 205
Db 941 EIAQVI-----APYEATSTQLSLTRQLIMIRKKTDSGWEGELQAKGRRR 987
QY 206 QTAWFPAPYLEEAPGQREGPSLGGSP-----QFCASRAYESSRADEL 251
Db 988 QIGWFPATYVKVL-----QGGRRNSGRNTPVSGSRIEMTEQILDKVIALYPYKAQNDDEL 1041
QY 252 SVPAGARVRVLETSDRGWMLCRYGDRAGLLPAVLLRP---EGLGALLSGT 298
Db 1042 SFDKDDIISVLGRDEPEWNGELNGLSGLFSPNSYVGFVTSKPKAKANGT 1091
RESULT 10
TI3055
dynamain associated protein isoform 160-2 - fruit fly (Drosophila melanogaster)
C:Species: Drosophila melanogaster
C:Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 09-Jul-2004
C:Accession: TI3055
R:Roos, J.; Kelly, R.B.

C;Species: Caenorhoditis eregans
C;date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C;Accession: T31504
R;McMurray, A.
submitted to the EMBL Data Library, October 1999
A;Reference number: Z21041
A;Accession: T31504
A;Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA
A:Residues: 1-1097 <W1>
A:Cross-references: UNIPROT:Q9U2T9; EMBL:AL117204; PIDN:CAB55138.1; CESP:Y116A8C.36
A:Experimental source: clone Y116A8C
C:Genetics:
A:Gene: CESP:Y116A8C.36
A:Introns: 50/3; 245/3; 411/3; 486/3; 697/1; 820/3; 926/2

Query Match 6.2%; Score 120.5; DB 2; Length 1097;
Best Local Similarity 19.9%; Pred. No. 1;
Matches 57; Conservative 35; Mismatches 107; Indels 87; Gaps 10;

QY 104 ATAERVARSPITITGFAPQPLDLEALPGSRVI-----LPTPEEPL 146
DB 727 AFVEIAAVPTPG--DPIQNPMPNTPSSVDQIGVKAARKABIAAMGUTEGAPPA 784
QY 147 SRAAGRLSIHSLEAQSRLCLPFCQTRDRPFOQAQESLDVLLRHPFGMWLVENEDRQ 206
DB 785 SSAPAAAAVIS-----QCIAQFQWRARNEEDLSPAKGDGTIEVLEKQEMK-WKGRNPAGE 837
QY 207 TAWFPAPYLE-----AAPQGREGSPSLGS-----SGPQ-- 236
DB 838 IGWFPKSYVKEGATTSTTTPTIVSPSKASAGAPGAAGAAQAYDVVPSDVTLOASETAPQOQ 897
QY 237 PCASRAVESSRADELSVPAGARVRLTSDRGWLCRYGDRAGLLPAVLLRPEGLGALL 295
DB 898 LTVIYDFEAVETTLALHVGDTIIVLEKNDE-WWKGRCNGREGIPPA----- 944
QY 296 SGTGPRGGDDPAGEARGPEPSQATAPPTVTR-PSPCAISQRC 340
DB 945 -----NYVEISVQAGDPTPTQAPTAAAPTIVLC 974

RESULT 14
T18552
saframycin Mx1 synthetase A - Myxococcus xanthus
C:Species: Myxococcus xanthus
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T18552
R:Pospiech, A.; Bietenhader, J.; Schupp, T.
Microbiology 142, 741-746, 1996
A:Title: Two multifunctional peptide synthetases and an O-methyltransferase are involved
A:Reference number: Z18967; MUID: 97090395; PMID: 8936303
A:Accession: T18552
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-2605 <POS>
A:Cross-references: UNIPROT:Q50858; EMBL:U24657; NID:g1171127; PID:g1171129; PIDN:AAC441
C:Genetics:
A:Gene: safA

C:Keywords: carrier protein
F:535-978/Domain: acetate-CoA ligase homology <ACL1>
F:997-1065/Domain: acyl carrier protein homology <ACP1>
F:1643-2091/Domain: acetate-CoA ligase homology <ACL2>
F:2110-2178/Domain: acyl carrier protein homology <ACP2>

Query Match 6.1%; Score 118.5; DB 2; Length 2605;
Best Local Similarity 24.9%; Pred. No. 3.7;
Matches 82; Conservative 33; Mismatches 105; Indels 109; Gaps 15;

QY 59 EAGLLRRSDRVLPKLLDAPLLGRVGRTS-----RGL-----ARLQLLETYSRRLATAE 107
DB 611 EAGLVGEGDANLEPVADAAQLAYVLYTSGSSGRPKGVWVSHGALANFLTITWAREPGLRAE 670
QY 108 RVARSPTITGFAPQPLDLEALPGSRVILPTPEEQLPSRA-AGRLSIHSLEAQSRLCL 166
DB 671 DVLAAVT-TFSPDIAALEIYLPVQGVWVATREQAADGRALSGVLARH----- 719
QY 167 QPFCTQTRDRPFOQAQESLDVLLRHPFGMWLVENEDRQTAWFPAPYLEAAPQCGREG 226
DB 720 -----GVTVMQATPATWRML-----ADA 737
QY 227 GPSLSSGGPQFCASRAYESSRADELSVPAGARV-RVLETSDRGWMLCR-----YGDAGLL 281

DB 738 GGAPGTGFTVLGCGEALPQDLADALTA-NGARVWNLVYGTETTVMWSCRKRLGAGDRVSLG 796
QY 282 PA-----VL-----LRPEGLGALLSGTFRGDDDPAGEARGF-----PEPSQA 319
DB 797 GALGNTSVHVLDPLRPPVPG--LAGELFIGG---SGVARGYWGRPSITAEFRFVDPFSA 851
QY 320 TAPPTVTRPSPGAIQSRCCCTVTRALE 348
DB 852 -----RPGARLYRTGDLVRRRVD 869

RESULT 15
A53800
mixed-lineage protein kinase (EC 2.7.1.1-) 3 - human
N:Alternate names: protein kinase PTKI; protein kinase SPRK
C:Species: Homo sapiens (man)
C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C:Accession: A53800; I58395
R:Gallo, K.A.; Mark, M.R.; Scadden, D.T.; Wang, Z.; Gu, Q.; Godowski, P.J.
J. Biol. Chem. 269, 15092-15100, 1994
A:Title: Identification and characterization of SPRK, a novel src-homology 3 domain-cont
A:Reference number: A53800; MUID:94253068; PMID:8195146
A:Accession: A53800
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-847 <GAL>
A:Cross-references: UNIPROT:Q16584; GB:U07747; NID:g464027; PIDN:AAA19647.1; PID:g464028
R:Ing, Y.L.; Leung, I.W.; Heng, H.H.; Tsui, L.C.; Lassar, N.J.
Oncogene 9, 1745-1750, 1994
A:Title: MLK-3: identification of a widely-expressed protein kinase bearing an SH3 domain
A:Reference number: I58395; MUID:94239754; PMID:8183572
A:Accession: I58395
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-847 <RES>
A:Cross-references: GB:L32976; NID:g488295; PIDN:AAA59859.1; PID:g488296
C:Genetics:
A:Gene: GDB:MLK3; PTKI; SPRK
A:Cross-references: GDB:I34755; OMIM:600050
A:Map position: 11q13.1-11q13.3
C:Superfamily: mixed-lineage protein kinase 3; protein kinase homology; SH3 homology
C:Keywords: ATP; leucine zipper; phosphotransferase; serine/threonine-specific protein ki
F:48-100/Domain: SH3 homology <SH32>
F:115-383/Domain: protein kinase homology <KIN>
F:123-131/Region: protein kinase ATP-binding motif
F:403-424/Region: leucine zipper motif
F:438-459/Region: leucine zipper motif
F:468-482/Region: basic

Query Match 6.1%; Score 118; DB 1; Length 847;
Best Local Similarity 34.3%; Pred. No. 1.1;
Matches 35; Conservative 9; Mismatches 38; Indels 20; Gaps 4;

QY 221 GGGP-EGGP-SLGSSGPQFCASRAYESSRADELSVPAGARVRL------TSDRGWMLCR 273
DB 28 GGGPESPKAAGYANPVWTALFDYFSGQDELARKGDRVEVLSRDAISGDEGWAQ 87
QY 274 YGDAGLLPAVLLRPEGLGALLSGTFRGDDDPAGEARGPPE 315
DB 88 VGGQVGIFP-----SNVSRGGGPPPCVASFOE 116

Search completed: May 28, 2005, 07:41:21
Job time : 43 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 28, 2005, 07:27:23 ; Search time 169 Seconds
(without alignments)
1124.150 Million cell updates/sec

Title: US-10-621-113-4
Perfect score: 1950
Sequence: 1 MAGPRYFVSQGAALVOIKR.....RRQGRPRGCVDSVPHTTEQ 371

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Uniprot_03.*
1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1950	100.0	371	2	Q8NFA3	Q8nfa3 homo sapien
2	1937.5	99.4	376	2	Q8NFA2	Q8nfa2 homo sapien
3	1934.5	99.2	370	2	Q96B73	Q96b73 homo sapien
4	1922	98.6	375	2	Q86YM1	Q86ym1 homo sapien
5	1219.5	62.5	349	2	Q8VCM2	Q8vcm2 mus musculus
6	1219.5	62.5	349	2	Q8BH41	Q8bh41 mus musculus
7	927	47.5	253	2	Q76EW0	Q76ew0 cavia porce
8	857.5	44.0	239	2	Q9D747	Q9d747 mus musculus
9	321	16.5	423	2	Q7T042	Q7t042 fugu rubrip
10	307.5	15.8	390	1	NCF1_HUMAN	P14598 homo sapien
11	307.5	15.7	389	2	Q811Y3	Q811y3 rattus norv
12	306.5	15.7	390	2	Q7TMH0	Q7tmh0 mus musculus
13	306	15.7	389	2	Q99M65	Q99m65 rattus norv
14	303.5	15.6	390	1	NCF1_MOUSE	Q90114 mus musculus
15	302	15.5	391	2	Q95M00	Q95m00 oryctolagus
16	300.5	15.4	392	1	NCF1_BOVIN	Q77774 bos taurus
17	299.5	15.4	392	2	Q95L71	Q95l71 bison bison
18	298	15.3	391	2	Q9N0E8	Q9n0e8 turlops tr
19	278	14.3	310	2	Q9JK56	Q9jk56 rattus norv
20	246.5	12.6	1124	2	O89032	O89032 mus musculus
21	245.5	12.6	563	2	O6DC19	O6dc19 brachydanio
22	240	12.3	471	2	Q8BIC6	Q8bic6 mus musculus
23	164	8.4	805	2	Q7S016	Q7s016 neurospora
24	163	8.4	174	2	Q76EV8	Q76ev8 cavia porce
25	161	8.3	989	2	O43302	O43302 homo sapien
26	157	8.1	1228	2	O6C7C0	O6c7c0 yarrowia li
27	151	7.7	1270	1	ITN1_XENLA	O42287 xenopus lae
28	148	7.6	1031	2	Q9H452	Q9h452 homo sapien
29	144.5	7.4	690	2	Q86WN1	Q86wn1 homo sapien
30	144.5	7.4	701	2	Q9NXX8	Q9nxx8 homo sapien
31	139.5	7.2	560	1	MIS_HUMAN	P03971 homo sapien

32	139	7.1	370	2	Q6UX75	Q6ux75 homo sapien
33	139	7.1	380	2	O55033	O55033 mus musculus
34	139	7.1	380	2	Q8BQ28	Q8bq28 mus musculus
35	139	7.1	380	2	Q6DBU6	Q6dbu6 brachydanio
36	138	7.1	410	2	Q24218	Q24218 drosophila
37	138	7.1	410	2	Q81PW2	Q81pw2 drosophila
38	138	7.1	548	2	Q9VPU1	Q9vpul drosophila
39	136.5	7.0	1168	1	MYSC_ACACA	P10569 acanthamoeb
40	136.5	7.0	1186	2	O61080	O61080 acanthamoeb
41	134	6.9	377	2	P79956	P79956 xenopus lae
42	134	6.9	380	1	NCK2_HUMAN	O43639 homo sapien
43	132	6.8	1183	2	Q7SH03	Q7sh03 neurospora
44	131	6.7	377	2	Q661L5	Q661l5 xenopus tro
45	131	6.7	403	2	Q803A2	Q803a2 brachydanio

ALIGNMENTS

RESULT 1
Q8NFA3 ID Q8NFA3 PRELIMINARY; PRT; 371 AA.
AC Q8NFA3;
DT 01-OCT-2002 (Tremblrel. 22, Created)
DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)
DT 05-JUL-2004 (Tremblrel. 27, Last annotation update)
DE Regulatory protein NOX01-beta (NADPH oxidase organizer 1).
GN Names=NOX01;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=14617635; DOI=10.1074/jbc.M305968200;
RA Cheng G., Lambeth J.D.;
RT "NOX01, regulation of lipid binding, localization, and activation of
RT Nox1 by the Phox homology (PX) domain."
RL J. Biol. Chem. 279:4737-4742(2004).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=22716264; PubMed=12716910; DOI=10.1074/jbc.M212856200;
RA Takeya R., Ueno N., Kami K., Taura M., Kohjima M., Izaki T., Nunoi H.,
RA Sumimoto H.;
RT "Novel human homologues of p47phox and p67phox participate in
RT activation of superoxide-producing NADPH oxidases."
RL J. Biol. Chem. 278:25234-25246(2003).
RN [3]
RP SEQUENCE FROM N.A.
RX PubMed=12473664; DOI=10.1074/jbc.C200613200;
RA Banfi B., Clark R.A., Steger K., Krause K.-H.;
RT "Two Novel Proteins Activate Superoxide Generation by the NADPH
RT Oxidase NOX1."
RL J. Biol. Chem. 278:3510-3513(2003).
CC - - SIMILARITY: Contains 2 SH3 domains.
DR EMBL; AF532984; AAM97926.1; - -
DR EMBL; AB097667; BAC76711.1; - -
DR EMBL; AF539796; AAN75141.1; - -
DR HSSP; P14598; 1KQ6.
DR GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR InterPro; IPR001683; PX.
DR InterPro; IPR001452; SH3.
DR Pfam; PF00787; PX; 1.
DR Pfam; PF00018; SH3; 1; 2.
DR ProDom; PD000066; SH3; 1.
DR SMART; SM00326; SH3; 2.
DR PROSITE; PS0195; PX; 1.
DR PROSITE; PS0002; SH3; 2.
KW SH3 domain.
SQ SEQUENCE 371 AA; 40796 MW; 39E13C3FF491797E CRC64;

Query Match 100.0%; Score 1950; DB 2; Length 371;
Best Local Similarity 100.0%; Pred. No. 4.3e-121;

```
Matches 371; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MAGPRYPVSVOGAALVQIKRLQTFAPSVRWSDGSDTFVRRSWDEFRQLKTKLKEFPVEA 60
Db 1 MAGPRYPVSVOGAALVQIKRLQTFAPSVRWSDGSDTFVRRSWDEFRQLKTKLKEFPVEA 60
QY 61 GLLRRSDRVLPKLLDAPLLGRVGRVTSRGLARLQLLETYSRRLLATAERVARSPFTTIGPFA 120
Db 61 GLLRRSDRVLPKLLDAPLLGRVGRVTSRGLARLQLLETYSRRLLATAERVARSPFTTIGPFA 120
QY 121 PQPLDLEPALPGSRVILPTPEEQPLSRAAGRLSIHSLAQSLRCLQPFCTQDTRDPFQ 180
Db 121 PQPLDLEPALPGSRVILPTPEEQPLSRAAGRLSIHSLAQSLRCLQPFCTQDTRDPFQ 180
QY 181 AQAQESLDVLLRHPSGWLWVENEEDQTAWFPAPYLEEAPGQREGGSLGSSGPFQFAS 240
Db 181 AQAQESLDVLLRHPSGWLWVENEEDQTAWFPAPYLEEAPGQREGGSLGSSGPFQFAS 240
QY 241 RAYESSRADELSPAGARVRVLETSDRGWMLCRYGDRAGLLPAVLLRPEGLGALLSGTGF 300
Db 241 RAYESSRADELSPAGARVRVLETSDRGWMLCRYGDRAGLLPAVLLRPEGLGALLSGTGF 300
QY 301 RGGDDPAGEARGFPPEPSQATAPPTVTRPSPGAIQSRCTVTRRALLERPRRQGRPGC 360
Db 301 RGGDDPAGEARGFPPEPSQATAPPTVTRPSPGAIQSRCTVTRRALLERPRRQGRPGC 360
QY 361 VDSVPHPTTEQ 371
Db 361 VDSVPHPTTEQ 371
RESULT 2
Q8NFA2 PRELIMINARY; PRT; 376 AA.
ID Q8NFA2
AC Q8NFA2;
DT 01-OCT-2002 (TremBLrel. 22, Created)
DT 01-OCT-2002 (TremBLrel. 22, Last sequence update)
DT 01-JUN-2003 (TremBLrel. 24, Last annotation update)
DE Regulatory protein NOXO1-gamma.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Lambeth J.D., Cheng G.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Contains 2 SH3 domains.
DR EMBL; AF532985; AAM97927.1; -.
DR HSSP; P14598; 1K06.
DR Genew; HGNC:19404; NOXO1.
DR GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR InterPro; IPR001683; PX.
DR InterPro; IPR001452; SH3.
DR Pfam; PF00018; SH3 1; 2.
DR ProDom; PD000066; SH3; 1.
DR SMART; SM00326; SH3; 2.
DR PROSITE; PS00195; PX; 1.
DR PROSITE; PS00002; SH3; 2.
KW SH3 domain.
SQ SEQUENCE 376 AA; 41252 MW; 2E9525A8BADD360 CRC64;
Query Match 99.4%; Score 1937.5; DB 2; Length 376;
Best Local Similarity 98.7%; Pred. No. 2.9e-120;
Matches 371; Conservative 0; Mismatches 0; Indels 5; Gaps 1;
QY 1 MAGPRYPVSVOGAALVQIKRLQTFAPSVRWSDGSDTFVRRSWDEFRQLKTKLKEFPVEA 60
Db 1 MAGPRYPVSVOGAALVQIKRLQTFAPSVRWSDGSDTFVRRSWDEFRQLKTKLKEFPVEA 60
QY 61 GLLRRSDRVLPKLLDAPLLGRVGRVTSRGLARLQLLETYSRRLLATAERVARSPFTTIGPFA 115
Db 61 GLLRRSDRVLPKLLDAPLLGRVGRVTSRGLARLQLLETYSRRLLATAERVARSPFTTIGPFA 120
```

```
QY 116 TGFAPOPLDLEPALPGSRVILPTPEEQPLSRAAGRLSIHSLAQSLRCLQPFCTQDTR 175
Db 121 TGFAPOPLDLEPALPGSRVILPTPEEQPLSRAAGRLSIHSLAQSLRCLQPFCTQDTR 180
QY 176 DRPFOAQAESLDVLLRHPSGWLWVENEEDQTAWFPAPYLEEAPGQREGGSLGSSGP 235
Db 181 DRPFOAQAESLDVLLRHPSGWLWVENEEDQTAWFPAPYLEEAPGQREGGSLGSSGP 240
QY 236 QFCASRAYESSRADELSPAGARVRVLETSDRGWMLCRYGDRAGLLPAVLLRPEGLGALL 295
Db 241 QFCASRAYESSRADELSPAGARVRVLETSDRGWMLCRYGDRAGLLPAVLLRPEGLGALL 300
QY 296 SCTGPRGDDPAGEARGFPPEPSQATAPPTVTRPSPGAIQSRCTVTRRALLERPRRQ 355
Db 301 SCTGPRGDDPAGEARGFPPEPSQATAPPTVTRPSPGAIQSRCTVTRRALLERPRRQ 360
QY 356 RPRGCVDSVPHPTTEQ 371
Db 361 RPRGCVDSVPHPTTEQ 376
RESULT 3
Q96B73 PRELIMINARY; PRT; 370 AA.
ID Q96B73
AC Q96B73;
DT 01-DEC-2001 (TremBLrel. 19, Created)
DT 01-DEC-2001 (TremBLrel. 19, Last sequence update)
DT 25-OCT-2004 (TremBLrel. 28, Last annotation update)
DE NADPH oxidase organizer 1, isoform a (NADPH oxidase regulatory
DE protein) (Regulatory protein NOXO1-alpha).
CN Name=NOXO1;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Colon;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Heiton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RN Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RP SEQUENCE FROM N.A.
RC TISSUE=Colon;
RA Strausberg R.;
RN Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
[3]
RP SEQUENCE FROM N.A.
RC TISSUE=Colon;
RX MEDLINE=22651106; PubMed=12657628; DOI=10.1074/jbc.M301289200;
RA Geiszt M., Leksstrom K., Witt J., Leto T.L.;
RT "Proteins homologous to p47phox and p67phox support superoxide
RT production by NAD(P)H oxidase 1 in colon epithelial cells."
RL J. Biol. Chem. 278:20006-20012(2003).
```

```

[4]
RN SEQUENCE FROM N.A.
RP Lambeth J.D., Cheng G.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Contains 2 SH3 domains.
DR EMBL; BC015917; AAI15917.1; -.
DR EMBL; AY255768; AAPI3479.1; -.
DR EMBL; AF532983; AAM97925.1; -.
DR HSSP; P14598; 1KQ6.
DR InterPro; IPR001452; SH3.
DR Pfam; PF00018; SH3_1; 2.
DR ProDom; PD000066; SH3; 1.
DR SMART; SM00326; SH3; 2.
DR PROSITE; PS0002; SH3; 2.
KW SH3 domain.
SQ SEQUENCE 370 AA; 40668 MW; 1EB8DCEB8BC50551F CRC64;

Query Match 99.2%; Score 1934.5; DB 2; Length 370;
Best Local Similarity 99.7%; Pred. No. 4.5e-120;
Matches 370; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 MAGPRYPVSVQGAALVQIKRLQTFAPSVRWSDGSDTFVRRSWDEFPQLKTLKETPVEA 60
Db 1 MAGPRYPVSVQGAALVQIKRLQTFAPSVRWSDGSDTFVRRSWDEFPQLKTLKETPVEA 59
QY 61 GLLRRSDRVLPKLLDAPLLGRVGTSGRLARLQLLETYSRRLLATAERVARSPPTITGFFA 120
Db 60 GLLRRSDRVLPKLLDAPLLGRVGTSGRLARLQLLETYSRRLLATAERVARSPPTITGFFA 119
QY 121 PQLDLEPALPPGSRVILTPPEQPLSRAAGRLSIHSLAQSLRCLQPFCTQTRDRPQ 180
Db 120 PQLDLEPALPPGSRVILTPPEQPLSRAAGRLSIHSLAQSLRCLQPFCTQTRDRPQ 179
QY 181 AQAQESLDVLLRHPSGWLVDNQTAWFPAPYLEEAPGQREGGSPISGSSGPOFCAS 240
Db 180 AQAQESLDVLLRHPSGWLVDNQTAWFPAPYLEEAPGQREGGSPISGSSGPOFCAS 239
QY 241 RAYESSRADELSPAGARVVRVLETSDRGWLCRYGDRAGLLPAVLLRPEGLGALLSGTGF 300
Db 240 RAYESSRADELSPAGARVVRVLETSDRGWLCRYGDRAGLLPAVLLRPEGLGALLSGTGF 299
QY 301 RGDDDPAGEARGPPEPSQATAPPTVTRPSPGAIQSRCTVTRRALERRRRPQGRPGC 360
Db 300 RGDDDPAGEARGPPEPSQATAPPTVTRPSPGAIQSRCTVTRRALERRRRPQGRPGC 359
QY 361 VDSVPHPTTEQ 371
Db 360 VDSVPHPTTEQ 370

RESULT 4
Q86YM1 PRELIMINARY; PRT; 375 AA.
ID Q86YM1
AC Q86YM1;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Regulatory protein NOX01-delta.
GN Name=NOX01;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]_
SEQUENCE FROM N.A.
RP Cheng G., Lambeth J.D.;
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Contains 2 SH3 domains.
DR EMBL; AV191359; AAO38665.1; -.
DR HSSP; P14598; 1KQ6.
DR InterPro; IPR001452; SH3.
DR Pfam; PF00018; SH3_1; 2.
DR ProDom; PD000066; SH3; 1.

[4]
RN SEQUENCE FROM N.A.
RP Lambeth J.D., Cheng G.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Contains 2 SH3 domains.
DR EMBL; BC015917; AAI15917.1; -.
DR EMBL; AY255768; AAPI3479.1; -.
DR EMBL; AF532983; AAM97925.1; -.
DR HSSP; P14598; 1KQ6.
DR InterPro; IPR001452; SH3.
DR Pfam; PF00018; SH3_1; 2.
DR ProDom; PD000066; SH3; 1.
DR SMART; SM00326; SH3; 2.
DR PROSITE; PS0002; SH3; 2.
KW SH3 domain.
SQ SEQUENCE 370 AA; 40668 MW; 1EB8DCEB8BC50551F CRC64;

Query Match 98.6%; Score 1922; DB 2; Length 375;
Best Local Similarity 98.4%; Pred. No. 3.1e-119;
Matches 370; Conservative 0; Mismatches 0; Indels 6; Gaps 2;

QY 1 MAGPRYPVSVQGAALVQIKRLQTFAPSVRWSDGSDTFVRRSWDEFPQLKTLKETPVEA 60
Db 1 MAGPRYPVSVQGAALVQIKRLQTFAPSVRWSDGSDTFVRRSWDEFPQLKTLKETPVEA 59
QY 61 GLLRRSDRVLPKLLDAPLLGRVGTSGRLARLQLLETYSRRLLATAERVARSPPTI 115
Db 60 GLLRRSDRVLPKLLDAPLLGRVGTSGRLARLQLLETYSRRLLATAERVARSPPTI 119
QY 116 TGFFAPQPLDLEPALPPGSRVILTPPEQPLSRAAGRLSIHSLAQSLRCLQPFCTQDTR 175
Db 120 TGFFAPQPLDLEPALPPGSRVILTPPEQPLSRAAGRLSIHSLAQSLRCLQPFCTQDTR 179
QY 176 DRPFOAQAQESLDVLLRHPSGWLVDNQTAWFPAPYLEEAPGQREGGSPISGSSGP 235
Db 180 DRPFOAQAQESLDVLLRHPSGWLVDNQTAWFPAPYLEEAPGQREGGSPISGSSGP 239
QY 236 QFCASRAYESSRADELSPAGARVVRVLETSDRGWLCRYGDRAGLLPAVLLRPEGLGALL 295
Db 240 QFCASRAYESSRADELSPAGARVVRVLETSDRGWLCRYGDRAGLLPAVLLRPEGLGALL 299
QY 296 SGTGFRGGDDPAGEARGPPEPSQATAPPTVTRPSPGAIQSRCTVTRRALERRRRPQ 355
Db 300 SGTGFRGGDDPAGEARGPPEPSQATAPPTVTRPSPGAIQSRCTVTRRALERRRRPQ 359
QY 356 RPRGCVDSVPHPPTTEQ 371
Db 360 RPRGCVDSVPHPPTTEQ 375

RESULT 5
Q8VCM2 PRELIMINARY; PRT; 349 AA.
ID Q8VCM2
AC Q8VCM2;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Noxol protein.
GN Name=Noxol;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]_
SEQUENCE FROM N.A.
RP STRAIN=FVB/N; TISSUE=Colon;
RC MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RX Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalek U., Smalish D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human

```

```

RT and mouse cDNA sequences.;
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N; TISSUE=Colon;
RA Strausberg R.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
CC -I- SIMILARITY: Contains 1 SH3 domain.
DR EMBL; BC019525; AAH19525.1; -.
DR HSSP; O89100; 10EB.
DR MGD; MGI:1919143; Noxol.
DR GO; GO:0016176; P:superoxide-generating NADPH oxidase activat. .; IDA.
DR GO; GO:0006801; P:superoxide metabolism; IDA.
DR Pfam; PF00787; PX; 1.
DR Pfam; PF00018; SH3_1; 2.
DR ProDom; PD000066; SH3; 1.
DR SMART; SM00312; PX; 1.
DR SMART; SM00326; SH3; 2.
DR PROSITE; PS01195; PX; 1.
DR PROSITE; PS00002; SH3; 1.
KW SH3 domain.
SQ SEQUENCE 349 AA; 38841 MW; 5008801247454101 CRC64;

Query Match 62.5%; Score 1219.5; DB 2; Length 349;
Best Local Similarity 67.4%; Pred. No. 8.1e-73;
Matches 242; Conservative 31; Mismatches 69; Indels 17; Gaps 4;

QY 1 MAGPRYPVSQGAALVQIKRQTFAFVSVDGSDTFVRRSWDEPRLQKTKLTFPVEA 60
Db |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
1 MASPRHPVSAHAVALQMDRLQTFAPFSCWSDNSDTFVRRSWDEPRLQKTKLTFPVEA 60
QY 61 GLLRRSDVLPKLDAPILGRVGRSGLARLQILETYSRLRLATAERVARSPITGFA 120
Db |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
61 GLLRRSEQVLPKLPAPLLTRGHTRGRLVRLRLDLYVQALLATSEHLRSSALHGFFV 120

121 PQLDLEPALPGSRVILPTPEEQPLSRAAGRLSIHSLEAOSLRCLQPCTQDTRDRPFQ 180
Db |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
121 PKPLDLEPMLPGSLVILPTPEE-PLSQPRSLDIHSLEAQSIPCVQPFHLDIRDRPFH 179

181 AOAGSLDVLRLHPSGWWLVNEDRQTAWFPAPYLEEAAPQGRGSPSLGSSGPFQFAS 240
Db |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
180 TKAQILDILLRHPSGWWLVNEDKQVAFPAFYLEEVATCQGBSGIALQSGRQFCTT 239

241 RAYESSRADELSPVAGARVRLVLETSRGMWLCRYGDRAGLAPAVLLRPEGLALLSGTGF 300
QY |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
240 QAYEGSRSDLSVPSGARVHVLETSRGMWLCRYNGRTGLLPAMSLQPEGLGSLLRPGF 299

301 --RGDDDPAGARGPPEPSQATAPPTVPTRPSPGAIQSRCTVTRRAL-----ERRPR 352
Db |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
300 PDSAGADKVAEDRTI-----PPVVTPECMASIQSRCCSITRRALAQEQGTRVPR 349

RESULT 6
Q8BH41 PRELIMINARY; PRT; 349 AA.
AC Q8BH41;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE NADPH oxidase organizer 1 (Mus musculus 2 days neonate thymus thymic
DE cells cDNA. RIKEN full-length enriched library, clone:E430007K11
DE product:231034C04RIK PROTEIN (SNX28) homolog).
GN Name=Noxol;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/c;
RX PubMed=12473664; DOI=10.1074/jbc.C200613200;
RA Banfi B., Clark R.A., Steger K., Krause K.-H.;
RT "Two Novel Proteins Activate Superoxide Generation by the NADPH

```

```

RT Oxidase NOX1.";
RL J. Biol. Chem. 278:3510-3513 (2003).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=NOD; TISSUE=Thymus;
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Mech. Enzymol. 303:19-44(1999).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=NOD; TISSUE=Thymus;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=NOD; TISSUE=Thymus;
RA The FANTOM Consortium;
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=NOD; TISSUE=Thymus;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=NOD; TISSUE=Thymus;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagao K., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kitsuai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishino T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaki S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kita A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
RN [7]
RP SEQUENCE FROM N.A.
RC STRAIN=NOD; TISSUE=Thymus;
RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA Nishi K., Nomura K., Numataki R., Ohno M., Ohsato N., Okazaki Y.,
RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
RA Tagawa A., Takahashi F., Takaku-Akashira S., Takeda Y., Tanaka T.,
RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
CC -I- SIMILARITY: Contains 1 SH3 domain.
DR EMBL; AF539797; AA075142.1; -.
DR EMBL; AK088226; BAC40222.1; -.
DR HSSP; O89100; 10EB.
DR MGD; MGI:1919143; Noxol.
DR GO; GO:0016176; P:superoxide-generating NADPH oxidase activat. .; IDA.
DR GO; GO:0006801; P:superoxide metabolism; IDA.
DR InterPro; IPR001683; PX.
DR InterPro; IPR001452; SH3.
DR Pfam; PF00787; PX; 1.
DR Pfam; PF00018; SH3_1; 2.

```



```
DR ProDom; PD000066; SH3; 1.
DR SMART; SM00312; PX; 1.
DR SMART; SM00326; SH3; 2.
DR PROSITE; PS50195; PX; 1.
DR PROSITE; PS50002; SH3; 1.
KW SH3 domain.
SQ SEQUENCE 349 AA; 38827 MW; 500F1E1247454101 CRC64;

Query Match 62.5%; Score 1219.5; DB 2; Length 349;
Best Local Similarity 67.4%; Pred. No. 8.1e-73;
Matches 242; Conservative 31; Mismatches 69; Indels 17; Gaps 4;

QY 1 MASPRPVSVQGAALVQIKLQTFAFSVVRSDSDTTFVRSWDEFRQLTKETFPVEA 60
DQ ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 MASPRPVSAVAHALVQMDRLQTFAFSVVCSNDSNDSNDSNDSNDSNDSNDSNDSN 60
QY 61 GLRRSDVLPKLLDAPLGRVGRGRTSGLARLQLETSYRRLATAERVARSPITGFFA 120
DQ ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 GLRRSEQVLPKLPDAPLLTRGHTGRLVRLRLDTYVQALLATSHEILRSSALHGFFV 120
QY 121 PQLDLLEPALPPGSRVILPTPEQPSLRRAAGRLSIHLSLEAQSRLCLOPCTQDTRDPFQ 180
DQ ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 121 PKPLDLEPMLPPGSLVILPTPEE-PLSQPRGSLDIHLSLEAQSIPCVQPFHTLDIRPFFH 179
QY 181 AQAQESLDVLRHPSGWLWVENEDROTANFPAPYLEEAPCGQREGGPGSLGSGPGFCAS 240
DQ ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 180 TKAQETLIDILRHPSGWLWVENEDQVAVFPAPYLEEAVATCGQESGLALQSGRQFCCT 239
QY 241 RAYESSRADELSVPAGARVRLVLETSRQWGLCRYGORAGLLPAVLLRPEGLGALLSGTGF 300
DQ ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 240 QAYEGSRDSELSVPSGARVHLVLETSRQWGLCRYNGRTGLLPAMSLQPEGLGSLGRPGF 299
QY 301 --RGSDPAGAEAGFPPEPSQATAPPVTPRPSPGATQSRCCCTWTRRAL-----ERRPR 352
DQ ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 300 PDSAGADKVAEDRTI-----PPVPTRPCMSALQSRCCSCTRALQGEQGTGVRPR 349

RESULT 7
Q76EW0 PRELIMINARY; PRT; 253 AA.
AC Q76EW0;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE NADPH oxidase organizer 1 (Fragment).
GN Names=NOXO1;
OS Cavia porcellus (Guinea pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Caviidae; Cavia.
OX NCBI_TaxID=10141;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Hartley;
RX PubMed=14978110;
RA Kawahara T., Kuwano Y., Teshima-Kondo S., Takeya R., Sumimoto H.,
RA Kishi K., Tsunawaki S., Hirayama T., Rokutan K.;
RT "Role of nicotinamide adenine dinucleotide phosphate oxidase 1 in
RT oxidative burst response to toll-like receptor 5 signaling in large
RT intestinal epithelial cells";
RT J. Immunol. 172:3051-3058(2004).
CC -1- SIMILARITY: Contains 1 SH3 domain.
DR EMBL; AB105906; BAD11766.1; -.
DR GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR InterPro; IPR001683; PX.
DR InterPro; IPR001452; SH3.
DR Pfam; PF00787; PX; 1.
DR ProDom; PD000066; SH3; 1.
DR SMART; SM00326; SH3; 2.
DR PROSITE; PS50195; PX; 1.
DR PROSITE; PS50002; SH3; 1.
KW SH3 domain.
FT NON_TER 1 1
FT 253 253

SQ SEQUENCE 253 AA; 28536 MW; 5616C7C89BE73124 CRC64;

Query Match 47.5%; Score 927; DB 2; Length 253;
Best Local Similarity 71.8%; Pred. No. 1.2e-53;
Matches 186; Conservative 18; Mismatches 49; Indels 6; Gaps 3;

QY 30 WSDGSDTFVRRSWDEFRQLTKETFPVEAGLLRRSDRVLPKLLDAPLGRVGRVTSRGL 89
DQ ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 WSDGSDTFVCRSWDEFRQLTKENFPVEAGLLRRSDRLPKLPDAPLLSRGRTGRGL 60
QY 90 ARLOLETYSRRLATAERVARSPITGFPAPQPLDLEPALPPGSRVILPTPEQPSLR 149
DQ ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 VRURLNTYVQELSLANVQSWSPVLTSFPEPRNQDLESALPPSSMVILEVP-EKPSRP 119
QY 150 AGRLSIHSLEAQSRLCLOPCTQDTRDRPFOAQESLDVLRHPSGWLWVENEDROTAW 209
DQ ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 120 VNSLDIHSLEIQSLCCVHPFHTQDTCGRPHVQAQESLDVLRHPSGWLWVENEGQKAW 179
QY 210 FPAPYLEEAPCGQREGGPGSLGSGPGFCASRAYESSRADELSVPAGARVRLVLETSR 269
DQ ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 180 FPAPYL-EVAP-----EVGLMLQNSGTVCASRAYKGSRADELSVPAGARVRLVLETSR 234
QY 270 WLCRYGDRAGLLPAVLLRP 288
DQ ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 235 WLCRYCDQGLLPVLLQP 253

RESULT 8
Q9D747 PRELIMINARY; PRT; 239 AA.
AC Q9D747;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Mus musculus adult male tongue cDNA, RIKEN full-length enriched
DE library, clone:2310034C04 product:2310034C04RIK PROTEIN (SNX28)
DE homology, clone:2310034C04 product:2310034C04RIK PROTEIN (SNX28)
GN Name=Noxol; Synonym=Snx28;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; Tissue=Tongue;
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RT Meth. Enzymol. 303:19-44(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; Tissue=Tongue;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; Tissue=Tongue;
RA The FANTOM Consortium;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; Tissue=Tongue;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RN [5]
RP Genome Res. 10:1617-1630(2000).
```


ID NCF1_HUMAN STANDARD; PRT; 390 AA.
 AC P14598; O43842; Q9BU90; Q9BX17; Q9BX18; Q9UDV9; Q9UMU2;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-MAY-1992 (Rel. 22, Last sequence update)
 DT 25-OCT-2004 (Rel. 45, Last annotation update)
 DE Neutrophil cytosol factor 1 (NCF-1) (Neutrophil NADPH oxidase factor
 DE 1) (47 kDa neutrophil oxidase factor) (p47-phox) (47 kDa
 DE autosomal chronic granulomatous disease protein) (NFXO2).
 GN Name=NCF1;
 OS Homo sapiens (Human).
 OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89332501; PubMed=2547247;
 RA Lomax K.J., Leto T.L., Nunoi H., Gallin J.I., Malech H.L.;
 RT "Recombinant 47-kilodalton cytosol factor restores NADPH oxidase in
 RT chronic granulomatous disease.";
 RL Science 245:409-412(1989).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89386707; PubMed=2550933;
 RA Volpp B.D., Nauseef W.M., Clark R.A.;
 RT "Cloning of the cDNA and functional expression of the 47-kilodalton
 RT cytosolic component of human neutrophil respiratory burst oxidase.";
 RL Proc. Natl. Acad. Sci. U.S.A. 86:7195-7199(1989).
 RN [3]
 RP REVISIONS.
 RA Volpp B.D., Nauseef W.M., Clark R.A.;
 RL Proc. Natl. Acad. Sci. U.S.A. 86:9563-9563(1989).
 RN [4]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90377229; PubMed=2398896;
 RA Rodaway A.R.F., Teahan C.G., Casimir C.M., Segal A.W., Bentley D.L.;
 RT "Characterization of the 47-kilodalton autosomal chronic granulomatous
 RT disease protein: tissue-specific expression and transcriptional
 RT control by retinoic acid.";
 RL Mol. Cell. Biol. 10:5388-5396(1990).
 RN [5]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97474758; PubMed=9329953;
 RA Gorlach A., Lee P.L., Roessler J., Hopkins P.J., Christensen B.,
 RA Green E.D., Chanock S.J., Curnutte J.T.;
 RT "A p47-phox pseudogene carries the most common mutation causing p47-
 RT phox-deficient chronic granulomatous disease.";
 RL J. Clin. Invest. 100:1907-1918(1997).
 RN [6]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=2038075; PubMed=10772875; DOI=10.1006/bcmd.2000.0274;
 RA Chanock S.J., Roessler J., Zhan S., Hopkins P., Lee P., Barrett D.T.,
 RA Christensen B.L., Curnutte J.T., Goerlach A.;
 RT "Genomic structure of the human p47-phox (NCF1) gene.";
 RL Blood Cells Mol. Dis. 26:37-46(2000).
 RN [7]
 RP SEQUENCE FROM N.A., AND VARIANT GLU-258.
 RC TISSUE=Umbilical vein;
 RX MEDLINE=21606173; PubMed=11740866; DOI=10.1006/excr.2001.5404;
 RA Gu Y., Xu Y.C., Wu R.P., Souza R.F., Nwariaku F.E., Terada L.S.;
 RT "TNFalpha activates c-jun amino terminal kinase through p47(phox).";
 RL Exp. Cell Res. 272:62-74(2002).
 RN [8]
 RP SEQUENCE FROM N.A., AND VARIANTS GLY-99 AND ASN-166.
 RX MEDLINE=22737999; PubMed=12853948; DOI=10.1038/nature01782;
 RA Hillier L.W., Fulton R.S., Fulton L.A., Graves T.A., Pepin K.H.,
 RA Wagner-McPherson C., Layman D., Maas J., Jaeger S., Walker R.,
 RA Wyllie K., Sekhon M., Becker M.C., O'Laughlin M.D., Schaller M.E.,
 RA Fowell G.A., Delehaunty K.D., Miner T.L., Nash W.E., Cordes M., Du H.,
 RA Sun H., Edwards J., Bradshaw-Cordum H., Ali J., Andrews S., Isak A.,
 RA Vanbrunt A., Nguyen C., Du F., Lamar B., Courtney L., Kalicki J.,
 RA Ozeraky P., Bielicki L., Scott K., Holmes A., Harkins R., Harris A.,
 RA Strong C.M., Hou S., Tomlinson C., Dauphin-Kohlberg S.,
 RA Kozlowski-Reilly A., Leonard S., Rohlfing T., Rock S.M.,
 RA Tin-Wollam A.M., Abbott A., Minx P., Maupin R., Stromatt C.,
 RA Latrelle P., Miller N., Johnson D., Murray J., Woessner J.P.,
 RA Wendt M.C., Yang S.P., Schultz B.R., Wallis J.W., Spieth J.,
 RA Bieri T.A., Nelson J.O., Berkowicz N., Wohldmann P.E., Cook L.L.,
 RA Hickenbotham M.T., Eldred J., Williams D., Bedell J.A., Mardis E.R.,
 RA Clifton S.W., Chissole S.L., Marra M.A., Raymond C., Haugen E.,
 RA Gillett W., Zhou Y., James R., Phelps K., Iadamoto S., Bubbs K.,
 RA Simms E., Levy R., Clendenning J., Kaul R., Kent W.J., Furey T.S.,
 RA Baertsch R.A., Brent M.R., Keibler E., Flicek P., Bork P., Suyama M.,
 RA Bailey J.A., Portnoy M.E., Torrents D., Chinwalla A.T., Gish W.R.,
 RA Eddy S.R., McPherson J.D., Olson M.V., Eichler E.E., Green E.D.,
 RA Waterston R.H., Wilson R.K.;
 RT "The DNA sequence of human chromosome 7.";
 RL Nature 424:157-164(2003).
 RN [9]
 RP SEQUENCE FROM N.A., AND VARIANTS GLY-99 AND ASN-166.
 RC TISSUE=Lymph;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Donald M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udell T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skaleka U., Smalhus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [10]
 RP SEQUENCE OF 8-31 FROM N.A.
 RC TISSUE=Ovary;
 RX MEDLINE=96090243; PubMed=7581362;
 RA Harshman K., Bell R., Rosenthal J., Katcher H., Miki Y., Swenson J.,
 RA Gholami Z., Frye C., Ding W., Dayananth P., Eddington K., Norris F.H.,
 RA Bristow P.K., Phelps R., Hattier T., Stone S., Shaffer D., Bayer S.,
 RA Hussey C., Tran T., Lai M., Rostock P.R. Jr., Skolnick M.H.,
 RA Shattuck-Eidens D., Kamb A.;
 RT "Comparison of the positional cloning methods used to isolate the
 RT BRCA1 gene.";
 RL Hum. Mol. Genet. 4:1259-1265(1995).
 RN [11]
 RP STRUCTURE BY NMR OF 1-128.
 RX MEDLINE=21267481; PubMed=11373621; DOI=10.1038/88591;
 RA Hiroaki H., Ago T., Ito T., Sumimoto H., Kohda D.;
 RT "Solution structure of the PX domain, a target of the SH3 domain.";
 RL Nat. Struct. Biol. 8:526-530(2001).
 RN [12]
 RP VARIANT CGD GLN-42, AND VARIANT SER-262.
 RX MEDLINE=20575420; PubMed=11133775; DOI=10.1182/blood.v97.1.305;
 RA Noack D., Rae J., Cross A.R., Ellis B.A., Newburger P.E.,
 RA Curnutte J.T., Heyworth P.G.;
 RT "Autosomal recessive chronic granulomatous disease caused by defects
 RT arising in the NCF1 pseudogenes.";
 RL Blood 97:305-311(2001).
 CC -!- FUNCTION: NCF2, NCF1, and a membrane bound cytochrome b558 are
 CC required for activation of the latent NADPH oxidase (necessary for
 CC superoxide production).
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -!- DISEASE: Defects in NCF1 are the cause of autosomal cytochrome-b-
 CC positive chronic granulomatous disease type 1 (CGD) [MIM:233700];
 CC an autosomal recessive form.
 CC -!- SIMILARITY: Contains 1 phox homology (PX) domain.


```

ID Q9NM65 PRELIMINARY; PRT; 389 AA.
AC Q9NM65;
DT 01-JUN-2001 (TREMELrel. 17, Created)
DT 01-JUN-2001 (TREMELrel. 17, Last sequence update)
DT 05-JUL-2004 (TREMELrel. 27, Last annotation update)
DE P47 phox (Neutrophil cytosolic factor 1).
GN Name=Ncf1;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley;
RA Tanabe M., Radmark O.P.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=DA;
RX MEDLINE=22397839; PubMed=12461526; DOI=10.1038/ng1058;
RA Olofsson P., Holmberg J., Tordsson J., Lu S., Akerstrom B.,
RA Holmdahl R.;
RT "Positional identification of Ncf1 as a gene that regulates arthritis
RT severity in rats.";
CC -!- SIMILARITY: Contains 2 SH3 domains.
DR EMBL; AY029167; AAK31797.1; -.
DR EMBL; AF547392; AAO32680.1; -.
DR HSSP; P14598; 1UEC.
DR GO; GO:0006118; P:electron transport; IEA.
DR GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR InterPro; IPR001655; P47PHOX.
DR InterPro; IPR001683; PX.
DR InterPro; IPR001452; SH3.
DR Pfam; PF00787; PX; 1.
DR Pfam; PF00018; SH3.1; 2.
DR PRINTS; PR00498; P47PHOX.
DR PRINTS; PR00452; SH3DOMAIN.
DR ProDom; PD000066; SH3; 2.
DR SMART; SM00312; PX; 1.
DR SMART; SM00326; SH3; 2.
DR PROSITE; PS0195; PX; 1.
DR PROSITE; PS00002; SH3; 2.
KW SH3 domain.
SQ SEQUENCE 389 AA; 44742 MW; BFC75842E53E68A4 CRC64;

Query Match
Best Local Similarity 27.1%; Score 306; DB 2; Length 389;
Matches 101; Conservative 55; Mismatches 155; Indels 62; Gaps 14;

QY 24 FAFSVRWSDGSDTFVRRSWDEFRQLKTKLTFPVEAGLLRRSDRVLPKLLDAPLLGRVG 83
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
26 YMFVVKQDLSKVVYKFTTEIYEPHKMLKEMFPIEAGEIHTENRVIHP-LPAPRWYDQG 84
QY 84 RTSRGLARLQLETSYRLLATAERVARSPITGFAFPQDLLEPALPGSRVILPTPEE 143
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
85 RAAE--SRQGLTTEYFNSLMGLPMKISRCPHLLNFFKVRPDDLK--LNDVSQVKKP--- 137
QY 144 QPLSRAAGRLSHSLEA---QSLRCLQFCTQDTRDRPFQQAQESLDVLLRHPSGWML 199
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
138 TYLTAKDGKNNVADIMGPIILLQTYRAIDY--EKSKTEMTVATGDVVDVVEKSESGWVF 195
QY 200 VENEDROTAWFPAPYLEA-APQREGGSPISGGSPQFCASRAYESSRADELSVPAGAR 258
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
196 CQMKTGR-GWNPASYLEPLDSPEADEDPDNY--AGEPYVTIKAYAAVEDEVSLSGEA 252
QY 259 VRVLETSRGRWLCRYGDRAGLLPAVLRPEG-----LGA----- 293
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
253 IEVHKHLDGWVVRKGDITGYFPMYLQKAGEETIQARQIRSGAPPRSTIRNAOSI 312
QY 294 -----LLSGTGFRCGD-----DPAGEARGFPE----PSQATA-PPPTVTPTRPSPCA 334
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
313 HQRSKRLSQDTYRRNSVRFLOQRREPARPGQSPDSKDNFSTTPRAKQPAPVPPRSSDL 372
```

```

QY 335 IQSRCCCTVTRRAL 347
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
373 ILHRCSTSTRKL 385

RESULT 14
NCF1_MOUSE
ID NCF1_MOUSE STANDARD; PRT; 390 AA.
AC Q03014; Q070144; Q9J134;
DT 01-NOV-1995 (Rel. 32, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Neutrophil cytosol factor 1 (NCF-1) (Neutrophil NADPH oxidase factor
DE 1) (47 kDa neutrophil oxidase factor) (p47-phox) (NCF-47K).
GN Name=Ncf1;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Macrophage;
RX MEDLINE=94164697; PubMed=8119734;
RA Jackson S.H., Malech H.L., Kozak C.A., Lomax K.J., Gallin J.I.,
RA Holland S.M.;
RT "Cloning and functional expression of the mouse homologue of
RT p47phox.";
RN Immunogenetics 39:272-275 (1994).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Leukemia;
RX MEDLINE=98149672; PubMed=9490028;
RA Mizuki K., Kadomatsu K., Hata K., Ito T., Fan Q.-W., Kage Y.,
RA Fukumaki Y., Sakaki Y., Takeshige K., Sumimoto H.;
RT "Functional modules and expression of mouse p40(phox) and p67(phox),
RT SH3-domain-containing proteins involved in the phagocyte NADPH oxidase
RT complex.";
RN Eur. J. Biochem. 251:573-582 (1998).
RN [3]
RP SEQUENCE FROM N.A.
RA Green E.D.;
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: NCF2, NCF1, and a membrane bound cytochrome b558 are
CC required for activation of the latent NADPH oxidase (necessary for
CC superoxide production).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: Contains 1 phox homology (PX) domain.
CC -!- SIMILARITY: Contains 2 SH3 domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L11455; AAA50469.1; -.
DR EMBL; AB002663; BAA25649.1; -.
DR EMBL; AF267747; AAF90134.1; -.
DR HSSP; P14598; 1UEC.
DR MGD; MGI:97283; Ncf1.
DR GO; GO:0016175; F:superoxide-generating NADPH oxidase activity; IEA.
DR GO; GO:0008283; P:cell proliferation; IMP.
DR GO; GO:0016066; P:cellular defense response (sensu Vertebrata); IMP.
DR GO; GO:0006954; P:inflammatory response; IMP.
DR GO; GO:0006691; P:leukotriene metabolism; IMP.
DR GO; GO:0009617; P:response to bacteria; IMP.
DR InterPro; IPR001655; P47PHOX.
DR InterPro; IPR001683; PX.
DR InterPro; IPR001452; SH3.
DR Pfam; PF00787; PX; 1.
```

Search completed: May 28, 2005, 07:44:16
Job time : 173 secs

This Page Blank (uspto)